Geographic and Temporal Variation in Housing Filtering Rates

Liyi Liu, Doug McManus, and Elias Yannopoulos
Freddie Mac
March 2020

Abstract

In the field of Housing Economics, filtering is the process by which properties, as they age, depreciate in quality and hence price and thus tend to be purchased by lower-income households. This is the primary mechanism by which competitive markets supply low-income housing. While at the national level filtering is an important long-term source of lower-income housing, this research shows that filtering rates for owner-occupied properties vary considerably both across and within metropolitan statistical areas (MSAs). Notably, in some markets, properties "filter up" to higher-income households. This paper contributes to our understanding of filtering by demonstrating the heterogeneity of filtering rates. The analysis finds strong geographic and temporal variation in filtering rates.

JEL R21, R31, R38

Any views expressed are solely those of the authors and do not necessarily represent the opinions of Freddie Mac or its Board of Directors. We thank Stuart Rosenthal, Kadiri Karamon, Salim Furth, Nancy Morrison and participants at FHFA's Econ Summit, the Housing Statistics Users Group, and AEI's Housing Collaborative seminars for helpful suggestions. Comments are welcome. e-mail: liyi_liu@freddiemac.com; douglas_mcmanus@freddiemac.com; elias_yannopoulos@freddiemac.com.

Introduction

In the United States, low-income housing is primarily created through filtering. New homes are largely purchased by higher-income households, but over time these homes depreciate and are purchased by lower-income households. This research uses a repeat income model, building on Rosenthal (2014), to estimate filtering rates at geographically and temporally disaggregated levels for owner-occupied properties. It shows that filtering rates for owner-occupied, single-family housing vary widely across and within metropolitan statistical areas (MSA) and across time.

In this paper, properties are said to *filter down* by 1% if for each new household that moves into the property, their real income is lower by 1% per year on average than the prior owner. Conversely, properties are said to *filter up* by 1% if the new household's real income increases by 1% per year on average. Rosenthal (2014) found that as owner-occupied properties age and new households move into these properties, the average real household income declines by about 0.7% per year (about 30% over 50 years). The results of this paper show a wide range of filtering rates, from rapid downward filtering in Chicago and Detroit to upward filtering in Washington, DC and Los Angeles. After modifying the repeat-income model by imposing a linearity assumption on filtering rates, which allows filtering rates to be estimated for 180 MSAs, the analysis shows that these filtering rates can range from about a 1.61% annual rate of downward filtering for Topeka, Kansas to a 0.71% upward filtering rate for San Francisco, California.[1]

Economic modeling of filtering in housing markets has a long history. Early theoretical research is presented in Sweeney (1974), which laid the foundation for theoretical models of filtering through depreciation offset by the level of maintenance. In contrast to the analytic approach of Sweeney (1974), Ohls (1975), in another early work, used numerical methods to solve his model and simulations to analyze the effect of government programs. Brueckner (1980) extended the traditional circular city model to allow properties to move between high-income and low-income occupants. Braid (1984) modified the model and mathematical techniques used by Sweeney to study the impacts of additional government interventions, including rent and income subsidies. Bond and Coulson (1989) considered a model of filtering where the income of occupants create an externality for the neighborhood. Galster and Rothenberg (1991) developed a model where housing markets are segmented by quality and household movement is endogenous to the model. Arnott and Braid (1997) further developed the model of Sweeney (1974) to include a more realistic maintenance function that allows properties to be maintained indefinitely.

Empirical filtering research has demonstrated the importance of filtering in housing markets. Most of the empirical filtering literature did not examine individual units directly but looked for outcomes consistent with filtering. Brueckner (1977) and Phillips (1981) used data to demonstrate factors that determine a Census tract's eligibility for residential succession, such as average effective rent, and the property age distribution. Sands (1979) provided the transition matrix among different quality properties using Michigan data. Weicher and Thibodeau (1988) studied the relationship between new construction and change in occupancy of low-quality housing at the city level. In rental markets, Weicher, Eggers, and Moumen (2016) reported that 45% of the rental units that were affordable to very

---

[1] The highest estimate of upward filtering is for Midland, Texas. This result is driven by the fracking boom since 2005 and so is not representative of filtering in this market over the sample period.

low-income renters in the United States in 2013 had filtered down from owner-occupied or higher rent categories in 1985. Recent work from Harvard's Joint Center for Housing Studies (2015) confirmed that filtering has been the primary source for additions to the affordable rental stock. Rosenthal (2014) provided the first direct empirical evidence of filtering by estimating repeat-income models based on repeat purchases of the same house by different households, but these estimates were limited to national and Census region levels because of sample limitations.

This paper builds upon Rosenthal (2014) by extending the repeat income model to MSA-level estimates and the structural model to various sub samples. It explores the differences in filtering rates for owner-occupied properties across MSAs by applying the repeat income model detailed in Rosenthal (2014) as well as a simplified linear model. This paper estimates the structural model, in which filtering rates are a function of depreciation and house price growth, on our data and finds qualitatively similar results to those in Rosenthal (2014). These structural model estimates are stable even after stratifying the sample by housing supply elasticity,[2] as in Saiz (2010), and when stratified by period, except for recent periods in which the effects of property age are greatly reduced. An implication of this structural model is that the heterogeneity of house price appreciation across MSAs will translate to similar heterogeneity of filtering rates across MSAs.

This paper also examines differential filtering rates *within* markets, specifically looking at variation in filtering rates within MSAs. A non-parametric analysis shows that filtering rates are far from uniform within MSAs. Most of the MSAs examined contain regions with both positive and negative filtering rates. The paper finds that the variability within MSAs is substantially larger than the variability across MSAs. Such heterogeneity within an MSA implies that even in markets with average upward filtering rates, some areas are creating affordable supply through filtering.

The variability of filtering rates has implications for housing policies. In markets with substantial downward filtering rates, filtering will likely be a robust source of supply for low-income housing. In contrast, markets with near-zero filtering rates or in which properties filter up, are likely to experience shortages of affordable properties without the construction of new homes for low income occupants.

In our sample, MSAs with elastic housing supply have faster downward filtering than the national average whereas markets with inelastic housing supply have upwards filtering on average.[3] Based on our estimated structural model most of this difference can be explained by the higher house price growth on average in inelastic markets. Thus, to the extent that the barriers to new supply are caused by restrictive land use and zoning regulations—such as limiting construction to single-family properties, imposing building height limits, requiring minimum lot sizes, or subjecting development to discretionary approval processes—then perhaps relaxing these restrictions would be an effective strategy for increasing affordable housing.[4]

---

[2] Supply elasticity measures how responsive supply is to prices: formally, it captures the percentage change in supply in response to a percentage change in price.

[3] Using supply elasticity estimates from Saiz (2010), MSAs with elasticity less than 1 are referred to as inelastic; those between 1 and 2 are deemed mid-elastic; and those greater than 2 are considered elastic.

[4] Mast (2019) has shown that the cascade of vacancies created from new luxury units also creates affordable housing through the chain of households moving into vacated properties. Policies that increase housing supply elasticity promote housing affordability by expanding the supply of market-rate housing at all levels. A substantial literature examines policies that are specifically aimed at increasing low-income housing. See, for example,

The heterogeneity of filtering rates within an MSA also raises policy questions. For example, the downward filtering rate in the Atlanta, Georgia MSA is driven by localized downward filtering to the south, east, and west of the city center, which offsets the upward filtering rates north of the city center. Such heterogeneity in within-MSA filtering rates may be related to greater income segregation in an MSA.[5] It also suggests that filtering is creating some affordable housing within the market, although in many cases this housing may be distant from the city center. It should also be kept in mind that while the focus of this paper is on the variability of filtering rates for owner-occupied properties, a more complete analysis would include an examination of the variability of filtering in rental units.

Data

We estimate filtering rates using repeated sales data created from home purchase transactions for owner-occupied properties with mortgages funded by Freddie Mac. We include mortgages originated from 1993 to 2018 for 1-unit single-family, condo, and townhouse properties built after 1900. Reliable income data are available only for loans originated from 1993 forward. We only use sales data where the property address was successfully scrubbed and uniquely identified as an actual street address. This address exclusion removed about 1.2 million sales from the population of 16.2 million. The key variable of interest in this data is the qualifying monthly income of the borrowers for the mortgage, which we use as a measure of household income.[6] To account for potential data quality issues, we exclude transactions with a missing purchase price or a price that exceeds $2 million, as well as those with a very low loan-to-value ratio of less than 20%. We also exclude transactions with very low incomes of $1000 or less per month and loans with a front-end debt-to-income ratio greater than 1 as well as very high incomes of $99,999 or more per month. These exclusions reduce the sample to 14.1 million sales.

We use tax assessor data sourced from CoreLogic© and Black Knight Data & Analytics, LLC for data on the year the structure was built and supplement this with the construction year reported by the seller from the Freddie Mac mortgage data. After incorporating these data, we had around 13.5 million sales for which we could identify the year built from one of these three sources. The CoreLogic tax assessor data also contain an effective year built, which accounts for significant additions to the properties, which we use in the robustness analysis. We use Consumer Price Index (CPI) data from the Bureau of Labor Statistics to convert nominal income to real income, with 2018 as the base year. For properties with multiple transactions, we create pairs of consecutive transactions and compute the change in the log real income between each pair. Of the 13.5 million sales, 2.53 million sales were for properties with two or more sales from which we can make the repeat observations. We then rule out *tear-downs*: that is, for each pair, we require the year built to differ by no more than one year across the sales, to exclude those cases in which the building was torn down and a new structure built. We assume that properties where the year built differ by a single calendar year refer to construction of the same structure and use

---

Freeman and Schuetz (2017); Malpezzi (2002); Sinai and Waldfogel (2005); Eriksen (2009); and Eriksen and Rosenthal (2010).

[5] For impacts of income segregation, see, for example, Chetty, Hendren, and Katz (2016).

[6] In some cases, changes in family structure—for example, through marriage or divorce—might introduce an additional factor to the incomes measured.

the older of the two reported construction years to compute the age for both sales.[7] We also restrict the sample to observations with changes in log real income per year between -0.6 and 0.6, which imposes approximately a 1 percent trim at each extreme.[8] This results in a sample of 1,233,888 repeat observations for 1,126,328 unique properties.

We use Freddie Mac's House Price Index (FMHPI) as an instrument in estimating structural models of filtering. MSAs are defined using ZIP Code data for the 2013 definitions from the Department of Housing and Urban Development (HUD). Each ZIP Code is assigned to the MSA containing most of the area of the ZIP Code.[9] For housing supply elasticity at the MSA level, we use the estimates in Saiz (2010). For data on regulatory restrictions on housing supply, we use the Wharton Residential Land Use Regulation Index (see Gyourko, Saiz, and Summers 2008). The latitude and longitude coordinates for the central business district (CBD) of MSAs come from Holian and Kahn (2015). Data on MSA-level population and median income are from the U. S. Census Bureau via Moody's Analytics. For data on improvements, we use data on appraisals for loans submitted for delivery to Fannie Mae and Freddie Mac since 2012 from the Uniform Appraisal Dataset (UAD).

Descriptive statistics for this data set are summarized in Table 1. The average number of years between transactions is 8.26 years. The overwhelming number of cases (91.18%) are for a single repeat of a given home. A modest share (8.14%) represents cases where a property has sold two times. Over the sample of repeats, real household income has declined by 3%, on average, between transactions.

Finally, the features of these data impose several limitations on our analysis. First, the data are restricted to loans funded by Freddie Mac, so this sample is likely to underrepresent low-valued properties, where the Federal Housing Administration (FHA) and subprime lenders have a large market share, and high-valued properties where loan amounts are beyond the conforming loan limit. To address this limitation and to provide a robustness check on our core results, we also implement a Heckman selection correction, leveraging data from the National Mortgage Database (NMDB). The NMDB is a nationally representative 5 percent sample of closed-end, first-lien residential mortgages in the United States and is funded and managed by Federal Housing Finance Agency and the Consumer Financial Protection Bureau.

An additional limitation is that the measured income used in this analysis is based on that reported by potential buyers to qualify for the loan, and in some cases may underestimate the income if a household reports only enough income needed for qualification. Finally, filtering rates for rental properties can be very important in determining the total market filtering rate, as Rosenthal (2014) demonstrated. However, our data cover only owner-occupied properties.

Model

[7] For example, a home is constructed in 2000 and sold for the first time in 2001. Sometimes, the first sale is recorded with year built as 2000 and the second sale with year built as 2001.

[8] In levels, this trims data where income increases annually by more than 82% (that is, [exp(0.6) – 1 = 0.82]) or when income decreases annually by more than 45% (that is, [exp(-0.6) - 1 = -0.45]).

[9] The HUD tool is available at https://www.huduser.gov/portal/datasets/geotools.html

Four methods are used to create estimates of filtering rates in this paper: 1) a repeat income model; 2) a simplification of this model assuming a linear form for the log income index; 3) a spatial localization of annual income changes; and 4) a structural model.

The repeat income method of Rosenthal (2014) produces an index of the household income as a function of property age. This method's structure is closely related to the repeat sales model (Bailey, Muth, and Nourse 1963), but is applied to income as opposed to the house price of the two sales and is indexed by the age of the home rather than the sale date. Following Rosenthal's (2014) notation, the model is specified as follows. For each transaction, the income of the arriving occupant at age t can be written as:

$$Y_t = e^{\gamma_t} f(X_t; \beta_t), \tag{1}$$

where $f(X_t; \beta_t)$ is an unknown and potentially nonlinear function of the structural and neighborhood characteristics of the home (X) and their shadow prices (β). The γ_t for each t are the parameters of interest and reflect the index of the household income of the arriving occupant(s) as the home ages in years. Assuming that X and β are constant over time, then the difference of the log of income across two periods becomes

$$\log\left(\frac{Y_{t+\tau}}{Y_t}\right) = \gamma_{t+\tau} - \gamma_t + \omega_t, \tag{2}$$

where $f(X; \beta)$ drops out of the model when taking the difference, and ω is the random error term. For a sample of properties with multiple transactions at various ages, the estimated model using linear regression is

$$\log\left(\frac{Y_{t+\tau,i}}{Y_{t,i}}\right) = \sum_{r=1}^{T} \gamma_r D_{r,i} + \omega_{r,i,} \tag{3}$$

where D_r equals -1 for the age at the time of the first transaction; 1 for the age of the second transaction; and 0 otherwise for each pair of transactions $(i \in 1, \ldots, N)$. The γ parameters can be estimated using linear regression of the log change in income on the matrix D. Provided that X and β are constant over time, the γ parameters are the filtering index.[10]

For smaller MSAs, we typically do not have enough data to estimate the index in each period using the repeat income technique. To overcome this problem, we impose the restriction that the log index is linear in time. With this added assumption, the repeat income specification becomes

$$\log\left(\frac{Y_{t+\tau}}{Y_t}\right) = g\tau + \omega_t, \tag{4}$$

where g is the filtering rate. To estimate filtering rates at even more granular levels, such as at the ZIP Code level, we use a nonparametric method: local polynomial regression. The technique is applied to a transformed variable,

[10] The results are reported in levels, which involves taking the exponent and setting the base level at 100.

$$\frac{\log\left(\frac{Y_{t+\tau}}{Y_t}\right)}{\tau}, \tag{5}$$

as the dependent variable, which is an observation-level estimate of the filtering rate. The independent variables in this regression are the longitude and latitude coordinates for each home. The R implementation of local polynomial estimation, “np,” is used to estimate filtering rates at each ZIP Code centroid. This method fits a linear specification using data restricted to a local neighborhood surrounding each evaluation point, where observations closer to the evaluation point are weighted more heavily. To achieve higher resolution in areas with more data, a smaller bandwidth was used to estimate filtering rates in ZIP Codes closer to the city center. For more background on this technique, see Fan and Gijbels (1996) and Li and Racine (2007). For a description of the “np” package, see Hayfield and Racine (2008).

The final model estimated is a structural model incorporating the effects of changes in house prices. Rosenthal (2014) specifies a structural model of housing demand where housing is decomposed into homogenous quality-adjusted units. The sum of the units is denoted by h, and the quality-adjusted price of the units is denoted q, yielding housing demand

$$\log(h_{t,i}) = \theta_Y \log(Y_{t,i}) + \theta_q \log(q_{t,i}), \tag{6}$$

where the parameters θ_Y and θ_q are the income and price elasticities of the demand for housing, respectively.

From this housing demand model, the change in log income is derived by solving for $\log(Y)$ and differencing across transactions. In addition, housing is assumed to depreciate at a constant annual rate ($\log(h_{t+\tau,i}/h_{t,i}) = d\tau_i$), yielding:

$$\log\left(\frac{Y_{t+\tau,i}}{Y_{t,i}}\right) = \frac{d}{\theta_Y}\tau_i - \frac{\theta_q}{\theta_Y}\log\left(\frac{q_{t+\tau,i}}{q_{t,i}}\right) + \omega_{t,i}. \tag{7}$$

Equation (7) implies that the filtering rates depend on the drivers of housing demand and the rate at which housing depreciates. Estimating this equation involves controlling for the quality-adjusted house price inflation, $\log(q_{t+\tau,i}/q_{t,i})$. It would seem natural to use the change in the observed transaction price as a proxy, but this would generate downward-biased estimates of filtering rates if estimated using ordinary least squares. As in Rosenthal (2014), this issue is addressed in a two-stage least squares (2SLS) regression by using differences in MSA-level house price indexes to instrument for the actual change in the house price.

Results

Repeat Income Model

At the national level, Figure 1 shows an index of the average income of households in a given property as the property ages. The horizontal axis is the age of the property, starting at 1 for new construction sold in the same year as it was built. The vertical axis is the repeat income index level, with the base level for the index set to 100 for a new property. This figure shows an average real income reduction of approximately 16% over 40 years, implying an average downward filtering rate of about 0.42% per year.

Thus, our data generate national estimates that are broadly in line with but lower than Rosenthal (2014). Our estimated index also has a more pronounced "U" shape, with the index showing upward filtering starting at an age of about 60 years and rebounding to within 5% of the real income for a new property after reaching 90 years.[11] This suggests that the gains in affordability for owner-occupied properties through filtering over the first 60 years of a property's life are reversed over the next 40. This "U" shape could be caused by lower quality older homes transitioning to rental properties or being torn down, and so dropping out of our sample.

The next exercise considers the variation in filtering rates across MSAs using the same repeat transactions estimator. Figure 2 displays estimates for six major MSAs: Atlanta, Chicago, Detroit, Los Angeles, Minneapolis, and Washington, DC. The figure clearly demonstrates substantial heterogeneity in filtering rates across these MSAs. After 40 years, average real incomes increased by 12.8% for Washington, DC (implying an average annual increase of 0.30%) and by 14.1% for Los Angeles (implying an average annual increase of 0.33%). Thus, properties in these markets were filtering up to higher-income households as homes aged. It is not surprising that these markets are ones with affordable housing challenges. In contrast, Detroit and Chicago show rapid downward filtering rates. For Detroit, the income index level drops 35.1% over 40 years (implying a rate of filtering of -1.1% per year). For Chicago, the income index level drops by 23.3% over 40 years (implying a rate of filtering of -0.66% per year).

Linear Model Estimates

To further investigate the heterogeneity of filtering rates across MSAs, the repeat income model in equation (4) is estimated with the simplifying assumption of a log linear index. This allows filtering rates to be estimated for many more MSAs because only a single parameter needs to be estimated rather than a time-varying index. Panels (a) and (b) of Table 2 provide results for the MSAs with the fastest and slowest filtering rates, where the more negative the coefficient, the faster the downward filtering rate.[12] All the filtering rates reported are statistically significantly different from zero. Panel (a) displays the 10 fastest MSA downward filtering rates. Topeka, Kansas has the fastest downward filtering rate, at roughly -1.61% per year. These are markets in which filtering is expected to provide a robust source of affordable supply. In contrast, Panel (b) displays MSAs with the slowest filtering rates. For example, San Francisco, California has an annual filtering rate of 0.71%, meaning that properties filter up to higher-income households as they age. Not surprisingly, the cities listed in Panel (b) tend to be regions with affordability problems. These tend to be markets where higher house price growth, on average, dominates depreciation to decrease the affordability of properties as they age. However, our spatial analysis described in the next section shows that there are often areas within these MSAs where downward filtering still occurs.

We explore potential drivers of variation in filtering rates across MSAs by plotting the MSA filtering rates against some relevant variables for housing markets: house price appreciation, supply elasticity, and population growth. Panel (a) of Figure 3 plots filtering rates by long-term average house price appreciation (HPA) and shows that house price growth is correlated with filtering rates. Panel (b) plots MSA filtering rates against housing supply elasticity as derived in Saiz (2010). There is a correlation

[11] Above the age of 96, the real income index is not statistically significantly different from 100.
[12] A complete list of the filtering estimates for 180 MSAs is in Table A1 of the appendix.

between filtering and supply elasticity, with more elastic markets having the fastest downward filtering rates. Two MSAs stand out in this plot: Austin, Texas which has an elastic supply but positive filtering; and Fargo, North Dakota, which has a very elastic supply but near-zero filtering. Both markets have experienced booms in local industries, with periods of above-average house price growth reflecting changes in housing demand not captured by the supply elasticity. This along with our regression analysis described later suggest that most of the effect of supply elasticity on filtering is through house prices. Panel (c) shows that the correlation of filtering rates with the Wharton Residential Land Use Regulation Index (WRI) is similar to that of the Saiz (2010) supply elasticity. Note, Fargo and Austin do not stand out in the WRI plot as they do in the supply elasticity plot. Panel (d) shows there is little association of MSA population growth with filtering rates. Panel (e) shows a positive correlation of MSA filtering rate with the share of appraisals with improvements (kitchen and bathroom remodeled within five years of appraisal). This relationship is similar if only kitchen remodels, only bathroom remodels, or kitchen or bathroom remodels are used as our measure of improvements. Finally, Panel (f) shows that the average MSA filtering rate is positively correlated with the growth in median income in the MSA, but downward filtering occurs in many MSAs with positive median income growth.

We use repeat purchase appraisals on same property from the UAD for 1-unit single-family properties from 2012 through 2019 submitted for delivery to Freddie Mac to compare the condition transition rates. We compare the condition transition of properties in upward filtering MSAs against those in downward filtering MSAs. A total of 81,597 appraisal pairs are used for the estimates of the top 30 downward filtering MSAs and 65,279 for the estimates of the top 10 upward filtering MSAs. The property condition variable is classified as follows: C1 = new construction/never occupied; C2 = recently renovated/like new; C3 = well maintained/limited depreciation; and C4 = adequately maintained/needs minimal repairs.[13] Panel (a) of Table 3 reports condition transitions for downward filtering MSAs. As shown in the third row, of the properties that have condition C3 for the first appraisal, 17% deteriorate to C4 at the second appraisal. Also, from the next row, 29% of properties in condition C4 remain in this condition. In contrast, in Panel (b), only 10% of the C3 properties deteriorate to C4 and only 19% of properties in condition C4 remain in this condition. This table shows that properties in upward filtering MSAs are less likely to transition to lower condition and those in lower condition are more likely to transition to an improved condition.

Nonparametric Models

Variation in filtering rates within MSAs is observed by estimating a spatial model that evaluates the filtering rate at the ZIP Code level using a nonparametric estimator, local linear estimation. Panel (a) of Figure 4 represents filtering rates estimated at every ZIP Code centroid in the Washington, DC metro area. There are clear differential filtering rates even within the city of Washington, DC. The southeast corner of the city beyond Anacostia has negative filtering rates, whereas the rest of the city has positive filtering rates. Outside of the city of Washington, DC is a band of strongly positive filtering rates that runs north to south of the city. To the east, the region in Prince Georges county adjacent to the city has strong negative filtering rates.

Panel (b) represents ZIP Code level filtering rates for the Atlanta metro area. The Atlanta, Georgia MSA has a negative average filtering rate, but there is variability within the MSA. Panel (b) shows positive

[13] Properties with condition below C4 are not purchased by Freddie Mac and thus are not included in our sample.

filtering rates near and to the north of the city center, and negative filtering rates outside the city center to the east, west, and south.

Panel (c) represents ZIP Code level filtering rates for the Chicago, Illinois metro area. The Chicago MSA has a negative average filtering rate, but has positive filtering rates in a band running southwest of the city center, with negative filtering rates to the south and north of that area.

Taken together, the three panels of Figure 4 demonstrate that filtering is far from uniform across space.[14] These MSAs provide examples where some regions within the MSA are creating new affordable housing through filtering, while other areas are not, as older properties filter up to higher-income households.[15] Generally, looking across many MSAs, we find that within most MSAs, upward filtering occurs in the areas closest to the city center, and that most often the highest downward filtering occurs outside the city center. The average of within-MSA variation over ZIP Code-level filtering rate estimates is substantially higher than the variance across the MSA-level filtering rate estimates.[16]

In the next analysis, we explore the relationship between house prices and variation of filtering rates within the MSA. Specifically, we observe that the city center of MSAs tend to have higher house price appreciation than the surrounding areas. To illustrate this phenomenon, we create an average estimate of the within-city variation in house price growth across MSAs. Because the MSAs are different sizes, we normalize the latitude and longitude of each MSA to range from -1 to 1, with the central business district located at the origin, (0,0), after the normalization. We apply this normalization to the latitude and longitude of each ZIP Code center within an MSA and create a combined data set across the MSAs used for analysis. We restrict the sample to the 26 MSAs with the most repeat income observations in our sample and use these for within-MSA filtering and the HPA analysis.[17] For each MSA, we calculate the ZIP Code-level HPA deviation from the MSA HPA mean and the observation's filtering rate minus the MSA mean filtering rate.

With the combined data, we estimate spatial variation in house price growth deviations from the MSA mean using kernel regression.[18] House price growth is measured as a ZIP Code's average annual house price growth from 1993 through 2018 based on the FMHPI. Panel (a) of Figure 5 plots the local linear estimates of house price growth for the interior of the city up to half of the way to the edge of the city in either direction. On average, the highest rates of HPA within the MSAs are heavily concentrated in ZIP

[14] In an earlier version, we verified the statistical significance of the within-MSA variations in filtering rates using the nonparametric test of Hsiao, Li, and Racine (2007) and as implemented in the npcmstest function of the R package "np."

[15] Of course, these are averages. Actual filtering is a stochastic process, so that even in markets that have upward filtering on average, some affordable properties can be created.

[16] Recently, Baum-Snow and Han (2019) found substantial variation in housing supply elasticity within metro areas.

[17] The 26 MSAs are Atlanta, Georgia; Austin, Texas; Boston, Massachusetts; Charlotte, North Carolina; Chicago, Illinois; Cincinnati, Ohio; Columbus, Ohio; Dallas, Texas; Denver, Colorado; Detroit, Michigan; Houston, Texas; Indianapolis, Indiana; Kansas City, Missouri; Los Angeles, California; Miami, Florida; Minneapolis, Minnesota; Nashville, Tennessee; New York, New York; Philadelphia, Pennsylvania; Phoenix, Arizona; Portland, Oregon; Raleigh, North Carolina; St. Louis, Missouri; Seattle, Washington; Tampa, Florida; and Washington, DC.

[18] The kernel regression is used instead of local linear regression to allow for a weighting that avoids MSAs with many observations dominating the results. For house price growth, this weighting is inversely proportional to the number of ZIP Codes with house price indexes for house price growth. For filtering rates, this weighting is inversely proportional to the number of repeat income observations in the MSA.

Codes near the city center.[19] Given the importance of house price growth in the structural model, we would also expect our analogous estimate of filtering rates to be higher near the city center. For Panel (b), we perform a similar exercise on the observation filtering rates minus the MSA average by pooling the data from the 26 MSAs and find that above-average filtering rates are also concentrated near the city center.

Structural Model

The structural model of Rosenthal (2014) offers insights into the drivers of filtering variability. This structural model is estimated using two-stage least squares (2SLS) with results reported in Table 4. The first-stage regression projects the change in log house prices onto the instrument, which is the corresponding percent change in the Freddie Mac House Price Index (FMHPI) at the MSA level when available, or the change in FMHPI at the Census division when the property is not in an MSA, to be consistent with Rosenthal (2014).[20] It is important for the index being used as the instrument to be at a high enough level of aggregation to be uncorrelated with the property-level error term in equation (7). Focusing on the first column, labeled "All," the results from the first-stage regression are reported in the first two rows. The change in these index values is positively related to house prices, as is reflected in the 0.611 regression coefficient. As can be seen from the large Kleibergan-Paap F-statistics, this is a strong instrument.[21]

The next three columns of this provide estimates stratified by Saiz (2010) supply elasticity. The "elastic" category includes MSAs with a supply elasticity greater than 2; the "mid-elastic" category consists of elasticities between 1 and 2; and the "inelastic" category includes MSAs with elasticities less than 1. The final column provides regression results using MSAs for which we do not have an elasticity from Saiz (2010). The results from these columns show coefficients that are close to the total sample estimates, suggesting that supply elasticity does not affect impact filtering beyond the effect on house prices.

The second-stage estimates of the structural model are given in the bottom of the table. The dependent variable is the log change in income. The key coefficient of interest is on years between transactions, with a value of -1.28%. It shows the contribution to the annual filtering rate of one additional year of property age. There are moderate differences in this coefficient across elasticity groups, with the downward filtering rate 20% faster for the elastic group than the inelastic group.[22] This model also shows the dependence of filtering rates on house price changes. From the first column, a 1% decrease in house prices results in a 0.356% increase in the filtering rate. The inelastic group filtering rates are around 19% more sensitive to changes in house prices than those of the elastic group.[23]

The volume of housing supplied, the expectation of future house price growth, the stringency of credit supply, and labor market conditions all varied during the study period. In Table 5, we estimate the structural model for four specific periods: pre-boom (January 1975 to December 2001); boom (January

---

[19] This figure may be of independent interest as it shows that while the spatial pattern of the HPA for any particular MSA can be complex, the average pattern of HPA across MSAs possess a simple monocentric spatial structure.
[20] Percent change is calculated as the HPI at time of the second sale divided by the HPI in the year of the first sale.
[21] In addition to the Kleibergan-Paap weak IV test, the Anderson-Rubin and Stock-Wright robust tests for weak instruments also show that this is a strong instrument inference, as the tests results are significant at the 0.1% level.
[22] This difference is statistically significant at the 10% level as the Z-statistic is 1.69.
[23] However, this difference just fails to be statistically significant at the 10% level, as the Z-statistic is 1.62.

2002 to June 2006); bust (July 2006 to December 2011); and post-crisis (January 2012 to December 2018). In each case, the sample is restricted to pairs of purchase transactions where both purchases occurred within the same subperiod. The first column shows the full sample results as presented in Table 4. The second through fifth columns of Table 5 provide the results for the four subperiods. Focusing on the years between transactions coefficients for the second stage, these coefficients are negative and significant in the first three periods (before 2012). This coefficient for the pre-boom and bust periods is around 1.2%, and is slightly higher (1.66%) for the boom period. However, for the post-crisis period (after 2011), the coefficient drops to 0.0698% and is not statistically significant, suggesting that the filtering process may have undergone a structural change in these years.[24]

An alternative way of examining the stability of the structural specification over time is to partition the impacts of time between transactions by time interval within a single regression. Table 6 provides estimates from this specification. In these regressions, the "years between transactions" variable in the prior models is segmented into four variables that represent the number of years between transactions that fell within each period (pre-boom, boom, bust, post-crisis). What is most interesting is that the coefficients change from the pre-boom estimate of -1.72% per year and boom estimate of -3.01% but then are closer to zero in the bust and post-crisis periods, at -1.02% and -0.632%, respectively. This also suggests a potential structural change, in that older properties are not filtering downward as fast as they had after conditioning on house price growth.

Gentrification

Finally, we examine the relationship between filtering and gentrification. Gentrification is characterized as the influx of higher socioeconomic status residents and an increase in housing prices. Tracts subject to gentrification are identified using the definition given in the National Community Reinvestment Coalition report, *Shifting Neighborhoods: Gentrification and Cultural Displacement in American Cities* (NCRC 2019). This measure identifies a tract as "gentrification eligible" if it has a population of at least 500 and had a median household income and median home value below 40th percentile within its MSA in 2000 and experienced an increase in median home value greater than the 60^{th} percentile and an increase in its share of college-educated residents greater than the 60^{th} percentile in its MSA.[25] We identify gentrification by assessing changes at the census tract level using nationwide U.S. Census Bureau data normalized by the longitudinal tract database (LTDB) between 2000 and 2012.[26]

Table 7 presents the regression analysis controlling for gentrification. We identify whether a property is in a gentrification tract and create an indicator variable for each pair. The first column focuses on the effects of gentrification on the annual filtering rate, which is defined as the log real income change divided by holding period, estimated by ordinary least squares (OLS). As seen from the first specification holding the change in log house prices constant, locating in a gentrification tract increases the log real

[24] This lack of statistical significance is not an issue of weak power, as the parameter is tightly estimated.
[25] There is no generally accepted definition of a gentrifying tract. For example, alternative definitions can be found in Freeman (2005) and Brummet and Reed (2019). These definitions all require the tract income to be low relative to the MSA median and increases in both the share of college-educated residents and house prices or rents.
[26] The longitudinal tract database (LTDB) provides estimates using 2010 boundaries for a standard set of variables from 1970 through 2000. The dataset includes a wide range of other variables based on sample data: the one-in-six samples from the decennial Census in 1970–2000 and the sample data from the American Community Survey for 2008–12. See https://s4.ad.brown.edu/projects/diversity/Researcher/Bridging.htm.

income by 0.37% in the annual filtering rate. For example, if the filtering rate were -0.5% in non-gentrifying tracts, it would be -0.13% in gentrifying tracts. In this sense, at the national level, gentrification is associated with slower downward filtering rates.

The second and third columns explore the effects of gentrification on filtering in the structural model. The second column shows the results of 2SLS when the sample is restricted to properties in the identified gentrification tracts. The sample size reduces to 12,520, around 1% of total transaction pairs counts. The coefficient on years between transactions is -0.836%, around 65% of the magnitude of the full sample estimate from Table 4. In the third column, we include interaction of the gentrification indicator with years between transactions in the 2SLS equation. Including the gentrification interaction term does not materially change the coefficient for years between transactions. Holding house price change and years between transactions constant, being in a gentrification tract increase the log real income change by 0.352% annually. These results indicate that gentrification tends to influence a market toward upward filtering. This is intuitive because higher-income households are attracted to gentrification areas because of additional neighborhood quality and amenity benefits that are not necessarily incorporated in house price increases. Note that this relationship need not be causal, given that inclusion in a gentrification tract may induce selection bias because the increase in median income of a tract is part of the definition of gentrification.

Robustness

Correcting for Selection Bias

The sample is restricted to loans funded by Freddie Mac and is likely to be underrepresented for low-valued properties, where the Federal Housing Administration (FHA) and subprime lenders have a large market share, and for high-valued properties where loan amounts are beyond the conforming loan limit. To address possible selection bias, we implement an extension of the Heckman (1979) correction as adapted to repeat sales models (see Gatzlaff and Haurin 1997; Hwang and Quigley 2004; Zanola 2007). This method first estimates a model to predict the probability of a mortgage being funded by government-sponsored enterprises (GSEs) based on a set of loan and borrower characteristics. In this first step, the probit model is fit using the NMDB data. Formally, let GSE_{it} be an indicator variable that takes the value of 1 if loan i, originated in year t, is funded by a GSE, and let:

$$Prob(GSE_{it} = 1) = \Phi\left(\gamma_0 + \sum_{m=1}^{M} \gamma_m Z_{mit}\right), \tag{8}$$

where Φ is the standard normal distribution, Z_{mit} are the loan characteristics ($m = 1 \ldots M$) of loan i originated at time t, and γ is a set of parameters.[27]The first stage estimates are based on more than 4 million purchase observations in the NMDB for loan originated between 1995 and 2018 with summary statistics in Table A2 and results in Table A3 of the appendix.

In the second step, we use the parameter estimates from the probit model to forecast sampling probabilities for our Freddie Mac data. We restrict our repeated income sample to loans originated

[27] Loan characteristics includes FICO score, loan-to-value ratio, debt-to-income ratio, loan amount, numbers of borrowers, numbers of units, loan term, loan product type (fixed or not), whether the borrower is a first-time home buyer, and origination year.

since 1995 to match the range of the NMDB data. Since the repeat income model's error term is subject to possible selection bias in both transaction times, the correction takes the form of the difference of two estimated bias terms. The inverse Mill's ratio— $\lambda_{t,i} = \phi(\gamma Z_{it})/\Phi(\gamma Z_{it})$—is included as an independent variable in equations (4) and (7) to correct for the non-randomness of sample selection. The structural model equation (7) becomes:

$$\log\left(\frac{Y_{t+\tau,i}}{Y_{t,i}}\right) = \frac{d}{\theta_Y}\tau_i - \frac{\theta_q}{\theta_Y}\log\left(\frac{q_{t+\tau,i}}{q_{t,i}}\right) + \beta(\lambda_{t+\tau,i} - \lambda_{t,i}) + \omega_{t,i}. \tag{9}$$

The corrected 2SLS results are presented in Table 8. The first column shows the previous estimates and the second column shows the coefficients after including the Heckman correction term $\beta(\lambda_{t+\tau,i} - \lambda_{t,i})$. The sample size reduces to 1,051,519 by restricting loans to be originated since 1995. The contribution to the annual filtering rate of one additional year of property age is -1.24%, and all the coefficient estimates are very similar to the results in Table 5. Restricting to loans originated after 1995 does not alter the filtering rate estimation. There is only a modest change for all the coefficient estimates after the Heckman correction. The coefficient for differences in the Mill's ratio is nonzero and significant, and the Wald test for this coefficient is significant, suggesting the existence of selection bias. The estimated coefficient for property age is -1.16%, which is a slightly weaker effect compared to the first column. The results show that the selection issue does not qualitatively change the results.

A similar Heckman selection adjustment can be applied to the linear filtering model presented in equation (4), resulting in the following estimation equation:

$$\log\left(\frac{Y_{t+\tau}}{Y_t}\right) = g\tau + \beta(\lambda_{t+\tau,i} - \lambda_{t,i}) + \omega_t. \tag{10}$$

Figure 6 plots the estimated filtering rates with and without the Heckman selection adjustment. Each point in this figure represents an MSA, with the X-axis plotting the filtering estimate with the Heckman correction, and the Y-axis plotting the filtering estimate without this correction. Since most of the data points lie close to the 45-degree line, the figure shows that the selection bias does not qualitatively change the linear model results.

Error Structure

In repeat sales models for measuring house price levels, the model errors are often modeled by a linear relationship in the time between sales. Case and Shiller (1987) argued that changes in house prices include components whose variance increases with the interval of sales, implying heteroskedastic error terms. We estimated this specification for filtering models by regressing the squared residuals on the time between sales.

We estimate the relationship between residual variance and change in property age for the repeat income model residuals in this paper and find that the change in property age has only a modest impact on the residual variance. This suggests that idiosyncratic changes in income mostly occur when a property sale takes place and depend only secondarily on the change in property age. For the national level repeat income model, the estimated equation for the variance of an observed filtering rate is $V_i(\tau) = 0.25 + 0.006 * \tau$, where τ is the time between sales in years. For the structural model the estimated equation is $V_i(\tau) = 0.25 + 0.005 * \tau$. In the data, the time between sales ranges from 1 to 25 years with a mean of 8.2 years. For this range of values, the intercept dominates the variance. The linear model estimated at the MSA level gives a range of estimated error structures, listed in appendix Table A4, with intercepts ranging from 0.16 to 0.38 and linear coefficients ranging from -0.001 to 0.027. However, the width of this range is driven by a few outliers. Looking at the range from the 10^{th} to 90^{th} percentile, the intercept ranges from 0.22 to 0.31 and the linear coefficients range from 0.0018 to 0.0087. This suggests that most MSAs have an error structure broadly in line with the national-level estimates.

Effective Age

The age of the property at the time of sale is a key variable in our analysis. In the main analysis for this paper, property age is determined based on the year of construction of the current structure. While this method accounts for a completely new structure built on the same lot, it does not account for any renovations to the property since it was initially constructed. The CoreLogic data contains an effective year built variable that reflects substantial renovations to the structure based on information from appraisals. We created a new dataset of pairs where both sales have the same effective year built. This gives a slightly smaller dataset than used for the main analysis: 1,190,584 pairs for 1,089,387 properties. This data set yields very similar estimates at the national and MSA level to the main analysis.

Figure 7 contains the national-level repeat income model estimates for effective age and construction age. The two estimates are very close (within 1.25, with a base level of 100) for the first 75 years of property age and diverge a bit more for the very oldest properties, where fewer data are available and estimates are more volatile. Figure 8 shows a comparison of the MSA level filtering estimates from the linear model. Of the 180 MSAs estimated, 37 have the exact same estimate because the effective age is the same as the construction age for all properties in the sample. There are 99 MSAs with lower filtering estimates when using the effective age data and 44 with higher estimates using effective age. Together, the mean absolute difference of these estimates is 0.0003, with a maximum difference of 0.0016. This suggests there may be a slight upward bias in the linear estimates from properties that have had significant renovations.

Filtering Relative to Median Income

Could some variation in filtering estimates simply reflect variation in income growth across MSAs? In Figure 3 panel (f), there is a strong positive correlation between filtering and real income growth over the sample period. As incomes increase throughout an MSA, it is possible for properties to go to occupants of higher real income with income farther below the median income. We explore this issue by estimating the linear model discounting by MSA median income rather than by the CPI. Formally, the quarterly MSA median income is converted to an index with the first quarter of 1993 as the base and the index is used to discount the household income in the same way as CPI was used in the original analysis. Figure 9 contains a scatterplot of the initial filtering estimates versus the estimates relative to the MSA median income. The biggest differences between the two estimates is for Midland, Texas on the downside and Flint, Michigan on the upside. Midland had the highest filtering estimate in real income at 0.86% per year but relative to median income of Midland the rate was -0.66% per year. In contrast, in Flint the filtering estimate was -1.3% in real terms but at a rate of 0.27% relative to median income. In general, those MSAs with upward filtering relative to real income had a lower rate, even going negative, relative to median income while MSAs with downward filtering rates had less negative rates, some going positive, relative to median income. Table A5 in the appendix contains the complete list of estimates using both methods.

Conclusion

Private markets provide affordable housing primarily through a process in which, on average, homes filter down to lower-income households as they age. Rosenthal (2014) established that, nationally, filtering is a major source of housing supply for low-income households. This paper contributes to our understanding of filtering by demonstrating the heterogeneity of filtering rates across space and time. The analysis finds strong geographic variation in filtering rates and strong differences in average filtering rates over time. There is substantial variation in filtering rates within MSAs in addition to variation across MSAs. As a result, markets with on average downward filtering have areas where incomes are rising—usually near the central business district, and markets with on average upward filtering have areas where properties are becoming more affordable. The variability of filtering rates is largely explained by differences in house price appreciation within and across MSAs.

The influence of house price appreciation on filtering implies a role for policy makers to adopt policies that would increase the elasticity of supply, driving down prices and allowing filtering to increase the stock of available affordable housing. Generally, policies encouraging the creation of new housing supply will assist in directly easing demand pressures and reducing the rate of house price appreciation, allowing properties to filter downward. Beyond filtering, policies that increase housing supply elasticity also promote housing affordability by directly expanding the supply of new affordable housing.

References

Arnott, R. J., and R. M. Braid. 1997. “A Filtering Model with Steady-State Housing.” *Regional Science and Urban Economics* 27 (4–5): 515–46.

Bailey, M. J., R. F. Muth, and H. O. Nourse 1963. “A Regression Model for Real Estate Price Index Construction.” *Journal of the American Statistical Association* 58: 933–42.

Baum-Snow, N., and L. Han. 2019. “The Microgeography of Housing Supply.” Working paper, University of Toronto.

Bond, E. W., and N. E. Coulson. 1989. “Externalities, Filtering, and Neighborhood Change.” *Journal of Urban Economics* 26 (2): 231–49.

Braid, R. M. 1984.” The Effects of Government Housing Policies in a Vintage Filtering Model.” *Journal of Urban Economics* 16 (3): 272–96.

Brueckner, J. 1977. **“**The Determinants of Residential Succession.” *Journal of Urban Economics* 4: 45–59.

---------. 1980. “Residential Succession and Land-Use Dynamics in a Vintage Model of Urban Housing.” *Regional Science and Urban Economics* 10: 225–40.

Brummet, Q., and D. Reed. 2019. “The Effects of Gentrification on the Well-Being and Opportunity of Original Resident Adults and Children.” Working Paper WP 19-30, Federal Reserve Bank of Philadelphia.

Case, K.E. and R.J. Shiller. 1987. “Prices of Single Family Homes since 1970: New Indexes for Four Cities.” *New England Economic Review* Sept/Oct: 45-56.

Chetty R., N. Hendren, and L. Katz. 2016. “The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment.” *American Economic Review* 106 (4): 855–902.

Eriksen, M. D. 2009. “The Market Price of Low-Income Housing Tax Credits.” *Journal of Urban Economics* 66 (2): 141–49.

Eriksen, M. D., and S. S. Rosenthal. 2010. “Crowd-out Effects of Place-based Subsidized Rental Housing: New Evidence from the LIHTC Program.” *Journal of Public Economics* 94 (11): 953–66.

Fan J., and I. Gijbels. 1996. *Local Polynomial Modelling and its Applications*. Chapman & Hall.

Freeman, L. 2005. “Displacement or Succession? Residential Mobility in Gentrifying Neighborhoods.” *Urban Affairs Review* 40 (4): 463–91.

Freeman L., and J. Schuetz. 2017. “Producing Affordable Housing in Rising Markets: What Works?” *Cityscape* 19 (1): 217–36.

Galster, G., and J. Rosenberg. 1991. “Filtering in Urban Housing: A Geographical Analysis of a Quality-Segmented Market.” *Journal of Planning Education and Research* 11 (1): 37–50.

Gatzlaff, D. H., and D. R. Haurin. 1997. “Sample Selection Bias in Local House Value Indices.” *Journal of Urban Economics* 43 (2): 199–222.

Gyourko, J., A. Saiz, and A. A. Summers. 2008. “A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index.” *Urban Studies* 45 (3): 693–729.

Hayfield T., and J. S. Racine. 2008. “Nonparametric Econometrics: The np Package.” *Journal of Statistical Software* 27 (5): 1–32.

Heckman, J. 1979. “Sample Selectivity Bias as Specification Error.” *Econometrica* 47 (1): 153–61.

Holian, M. J., and M. E. Kahn. 2015. “Household Carbon Emissions from Driving and Center City Quality of Life.” *Ecological Economics* 116: 362–68.

Hsiao, C., Q. Li, and J. S. Racine. 2007. “A Consistent Model Specification Test with Mixed Discrete and Continuous Data.” *Journal of Econometrics* 140: 802–26.

Hwang, M., and J. M. Quigley. 2004. “Selectivity, Quality Adjustment and Mean Reversion in the Measurement of House Values.” *Journal of Real Estate, Finance and Economics* 28 (2/3): 161–78.

Joint Center for Housing Studies for Harvard. 2015. *America’s Rental Housing: Expanding Options for Diverse and Growing Dema*nd. Cambridge, MA: Harvard University.

Li, Q., and J. S. Racine. 2007. *Nonparametric Econometrics: Theory and Practice.* Princeton, NJ: Princeton University Press.

Malpezzi S. 2002. “Does the Low-Income Housing Tax Credit Increase the Supply of Affordable Housing?” *Journal of Housing Economics* 11 (4): 360–80.

Mast, E. 2019. “The Effect of New Market-Rate Housing Construction on the Low-Income Housing Market.” Upjohn Institute Working Paper 19-307, W. E. Upjohn Institute for Employment Research, Kalamazoo, MI. https://doi.org/10.17848/wp19-307.

NCRC (National Community Reinvestment Coalition). 2019. *Shifting Neighborhoods: Gentrification and Cultural Displacement in American Cities*. Washington, DC: National Community Reinvestment Coalition.

Ohls, J. C. 1975. **“**Public Policy toward Low-Income Housing and Filtering in Housing Markets.” *Journal of Urban Economics* 2 (2): 144–71.

Phillips, R. S. 1981. “A Note on the Determinants of Residential Succession.” *Journal of Urban Economics* 9 (1): 49–55.

Rosenthal, S. S. 2014. "Are Private Markets and Filtering a Viable Source of Low-Income Housing? Estimates from a 'Repeat Income' Model." *American Economic Review* 104 (2): 687–706.

Saiz, A. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics* 125 (3): 1253–96.

Sands, G. 1979. "A Model for the Evaluation of Filtering." *Growth and Change* 10 (4): 20–24.

Sinai, T., and J. Waldfogel. 2005. "Do Low-Income Housing Subsidies Increase the Occupied Housing Stock?" *Journal of Public Economics* 89 (11–12): 2137–64.

Sweeney, J. L. 1974. "A Commodity Hierarchy Model of the Rental Housing Market." *Journal of Urban Economics* 1 (3): 288–323.

Weicher, J., F. Eggers, and F. Moumen. 2016. "The Long-Term Dynamics of Affordable Rental Housing." Washington, DC: Hudson Institute.

Weicher J., and T. J. Thibodeau. 1988. "Filtering and Housing Markets: An Empirical Analysis." *Journal of Urban Economics* 23 (1): 21–40.

Zanola, R. 2007. "The Dynamics of Art Prices: The Selection Corrected Repeat-Sales Index." POLIS Working Paper 76, Institute of Public Policy and Public Choice.

Table 1. Summary statistics

Variable	Mean
Years between transactions	8.26
log change in real income	-0.03
Monthly income at time of transaction ($2018)	9299
Age of home at time of transaction (years)	23.32
House price at time of transaction ($)	235,929
log change in house price	0.27
Distribution of transaction pairs per home (percent)	
1 pair	91.18
2 pairs	8.14
3 pairs	0.63
4+ pairs	0.04
Number of homes	1,126,328
Repeat observations	1,233,888

Figure 1. Filtering index of owner-occupied properties

The index tracks the average level of occupant real income as the home ages, estimated using the repeat income model.

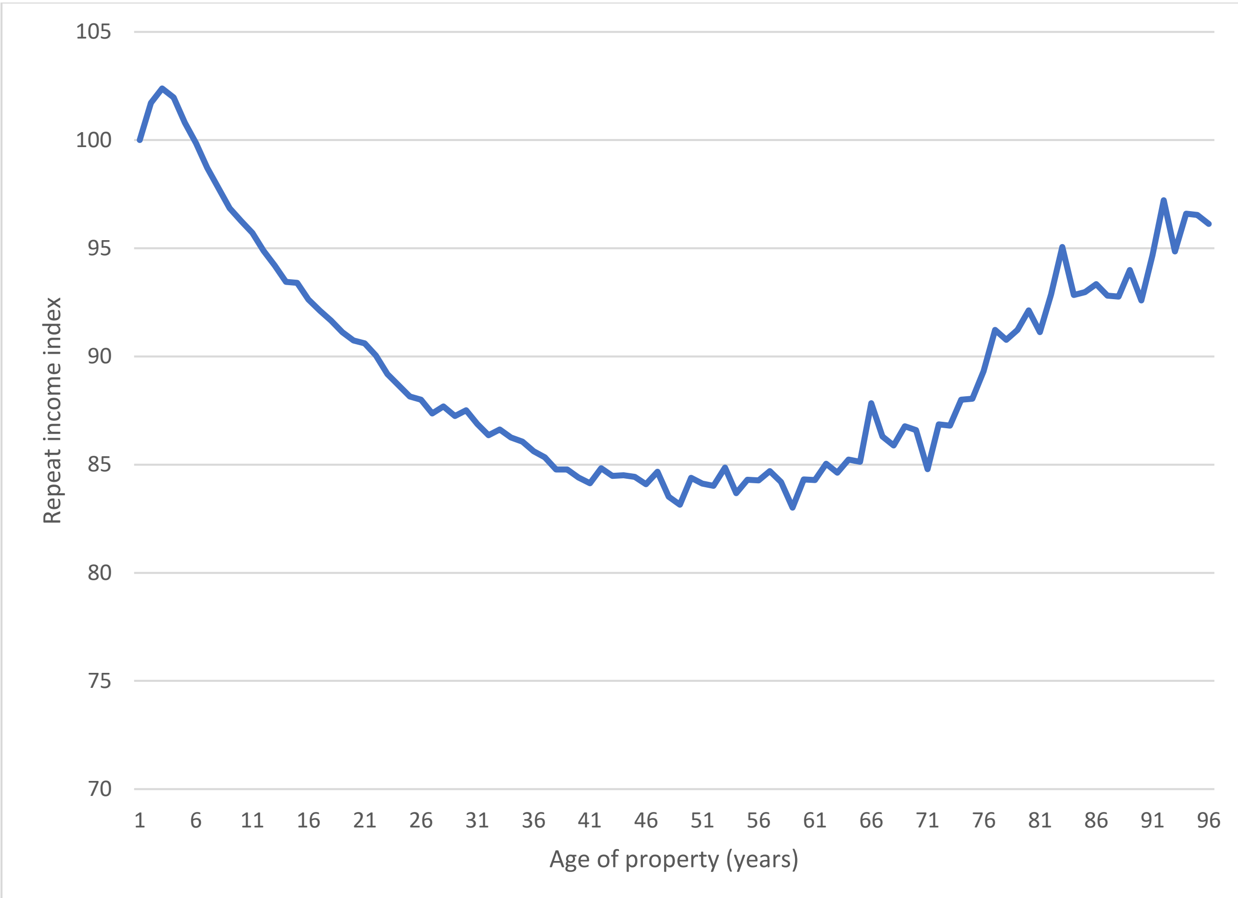


Figure 2. Filtering index by MSA

Filtering estimates using the repeat income model are quite heterogeneous at the MSA level. For example, Los Angeles, CA and Washington, DC show substantial increases in the real income of occupants as properties age, while, Detroit, MI and Chicago, IL show a real income decline over the same period.

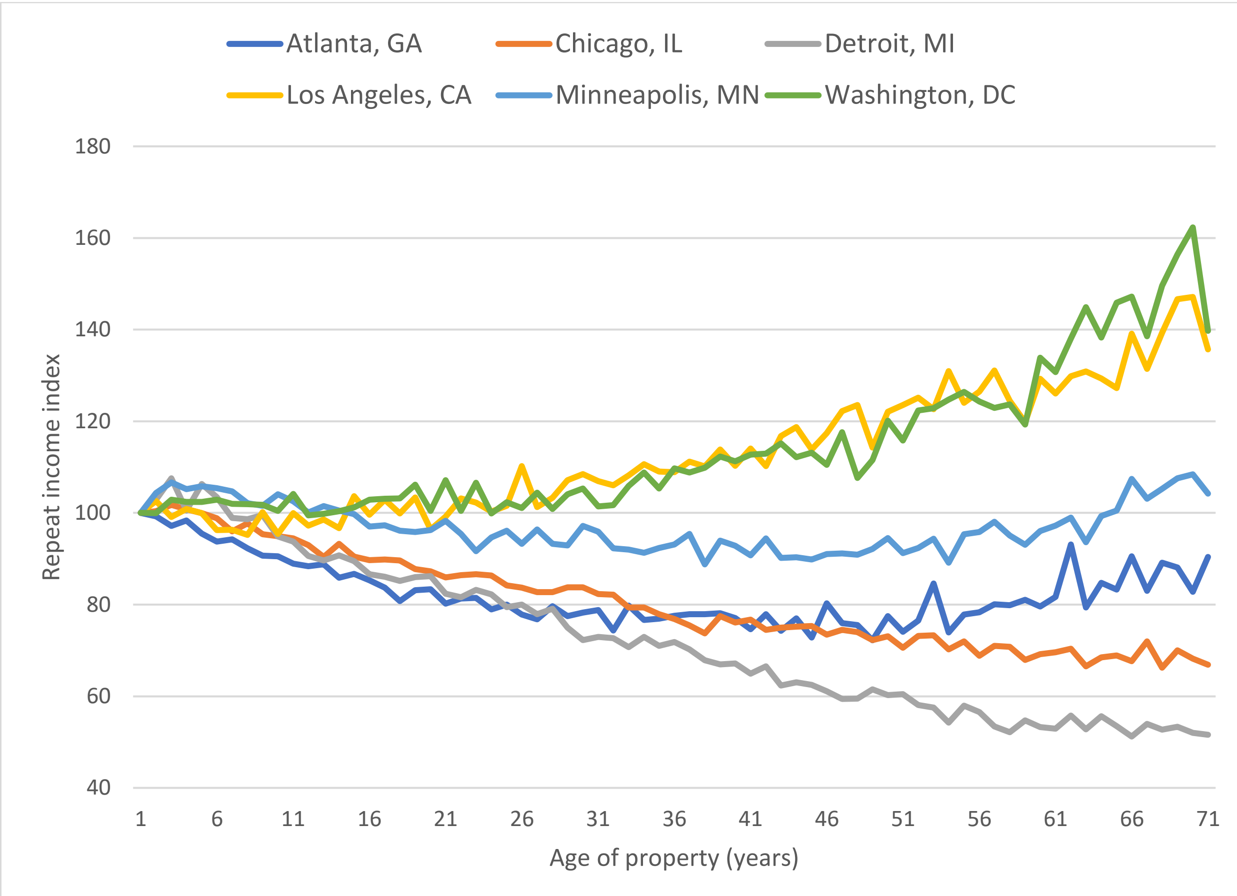


Table 2. Linear model of real change in income (log) by MSA

a. MSAs with largest reduction in occupant income

MSA	Annual Filtering (log)	Standard Error	40-Year (%)
Topeka, KS	-0.0161	0.0018	-47.38
Macon, GA	-0.0149	0.0020	-44.79
Jackson, MS	-0.0148	0.0018	-44.59
Fort Wayne, IN	-0.0145	0.0009	-43.94
Toledo, OH	-0.0135	0.0008	-41.66
Flint, MI	-0.0132	0.0015	-40.95
South Bend, IN	-0.0131	0.0013	-40.86
Myrtle Beach, SC	-0.0130	0.0021	-40.62
Spartanburg, SC	-0.0130	0.0018	-40.55
Greensboro, NC	-0.0125	0.0008	-39.32

b. MSAs with largest increase in occupant income

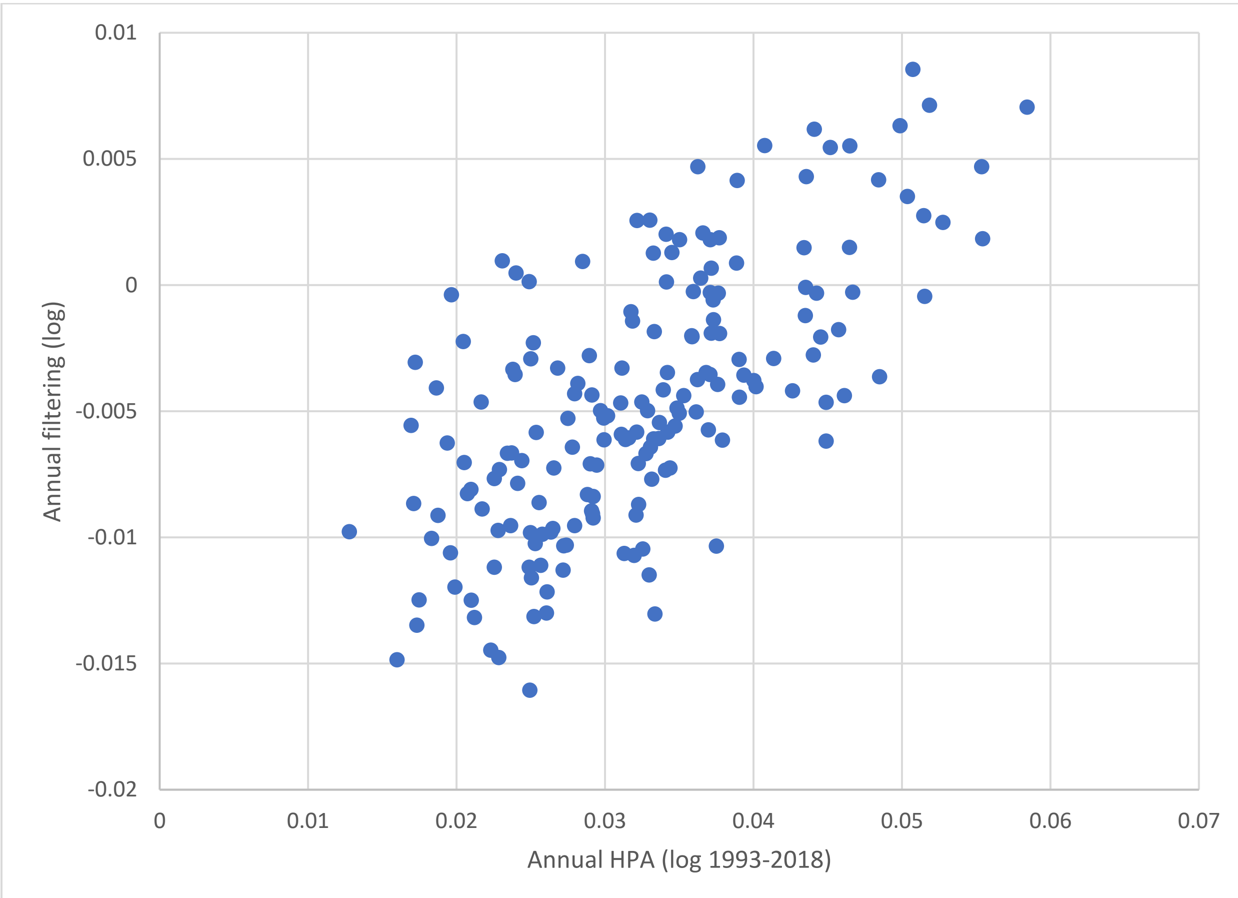
MSA	Annual Filtering (log)	Standard Error	40-Year (%)
Boulder, CO	0.0047	0.0009	20.64
Charlottesville, VA	0.0047	0.0018	20.68
Los Angeles, CA	0.0055	0.0004	24.41
San Diego, CA	0.0055	0.0006	24.71
Oxnard, CA	0.0055	0.0010	24.76
Santa Rosa, CA	0.0062	0.0013	28.04
Seattle, WA	0.0063	0.0004	28.76
San Jose, CA	0.0071	0.0008	32.58
San Francisco, CA	0.0071	0.0006	33.00
Midland, TX	0.0086	0.0015	40.78

Note: The filtering rate is estimated using equation (4), the linear model with no intercept. The 40-year percentage change is computed as $[\exp(40 * Filtering\ coefficient) - 1] * 100$. Only the first city of each MSA is listed.

Figure 3. MSA annual filtering (log)

a. By house price appreciation (HPA)

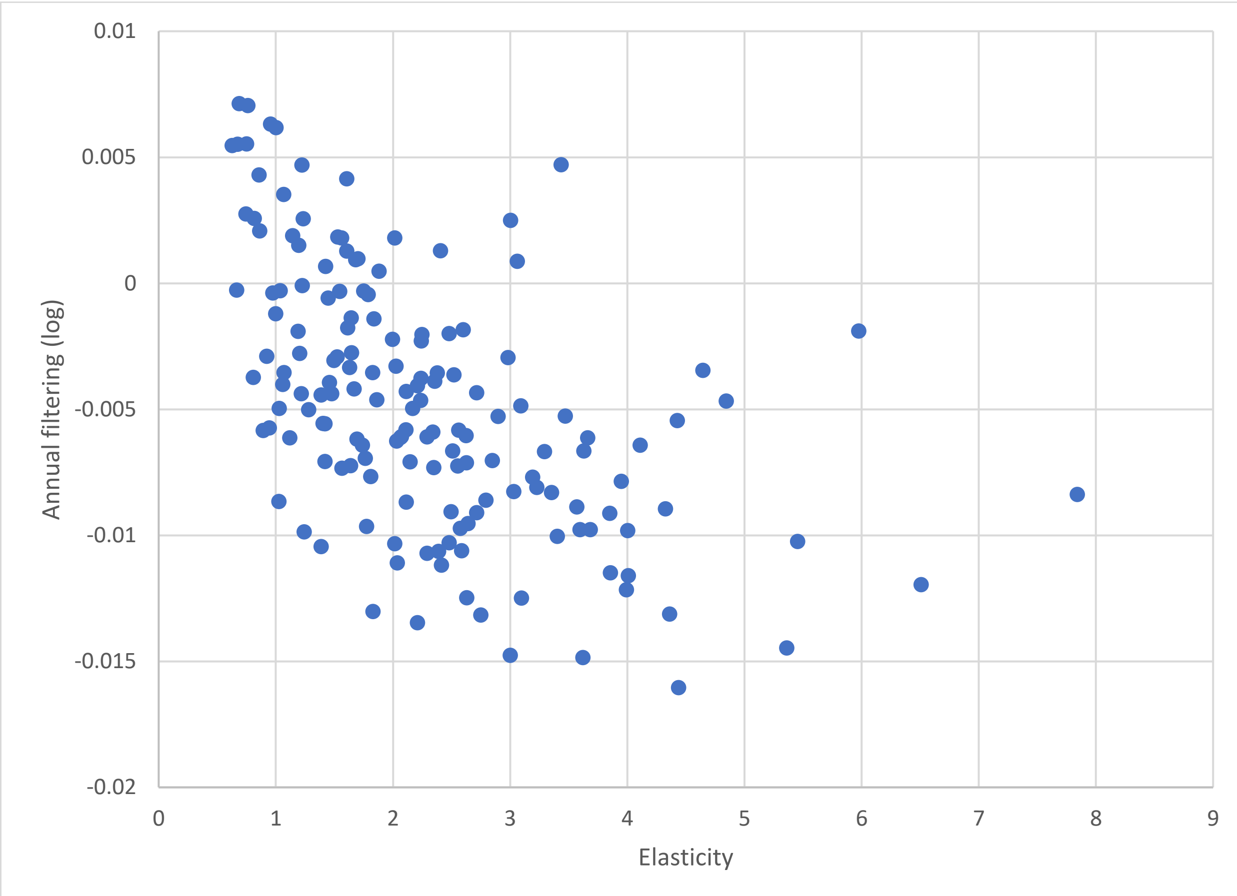
The relationship between long-term HPA and filtering rates across MSAs is positive, as expected.



Note: Log annual filtering rates are those reported in Table A1. Average annual house price appreciation is estimated from the log of the Freddie Mac house price index.

b. By supply elasticity

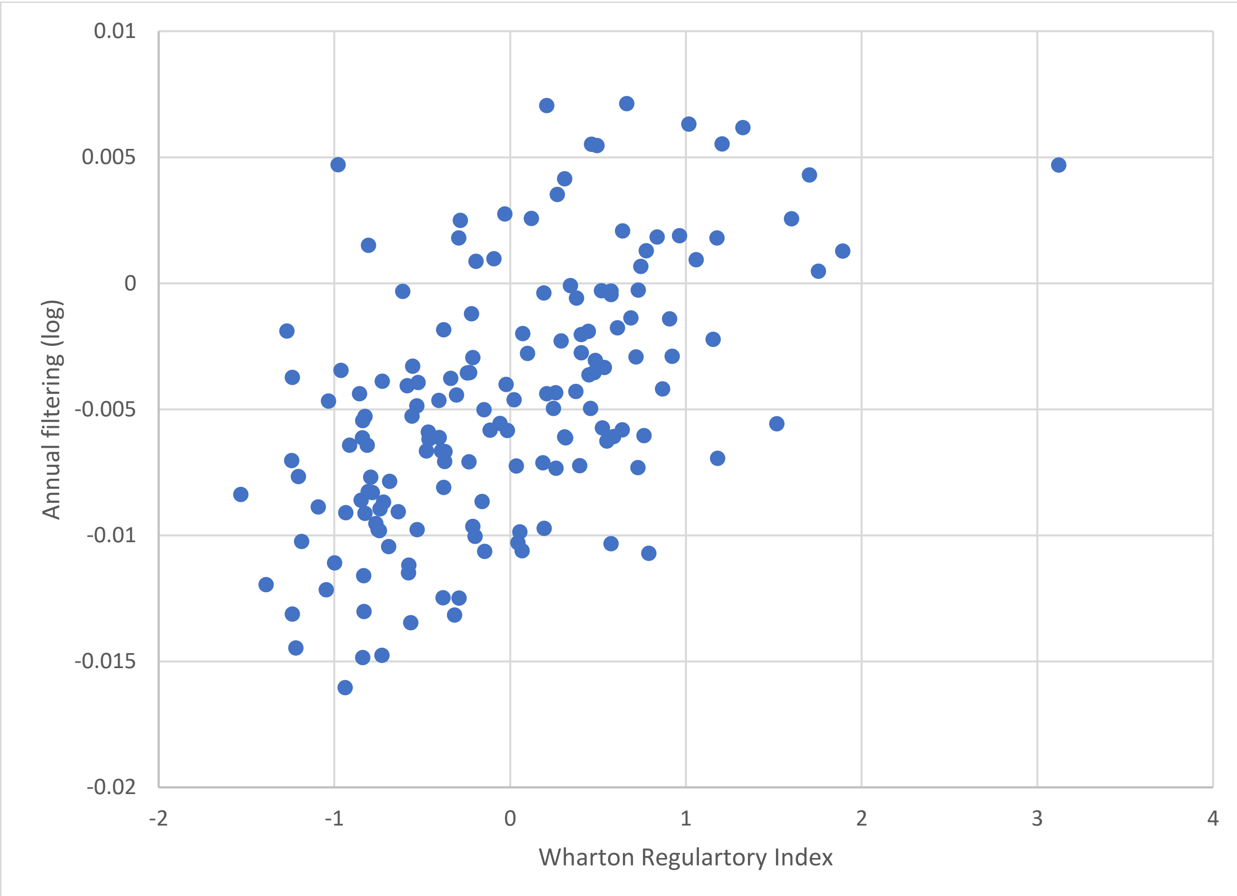
The relationship with supply elasticity and filtering is not as strong as with house prices.



Note: Log annual filtering rates are those reported in Table A1. Elasticity refers to housing supply elasticity measured by Saiz (2010).

c. By the Wharton Regulatory Index

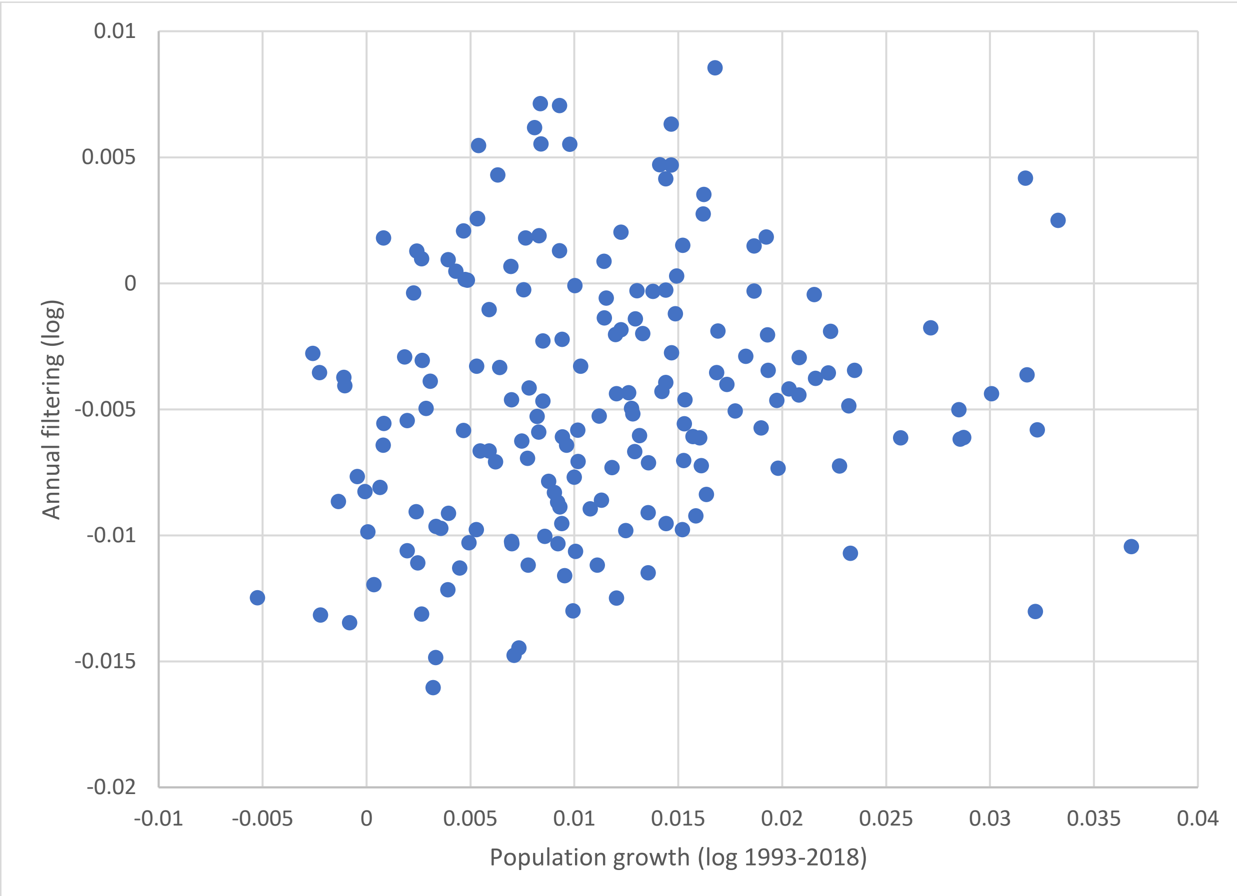
The relationship between the Wharton Regulatory Index and filtering is not as strong as with house prices.



Note: Log annual filtering rates are those reported in Table A1. The Wharton Regulatory Index is from Gyourko, Saiz, and Summers (2008).

d. By population growth

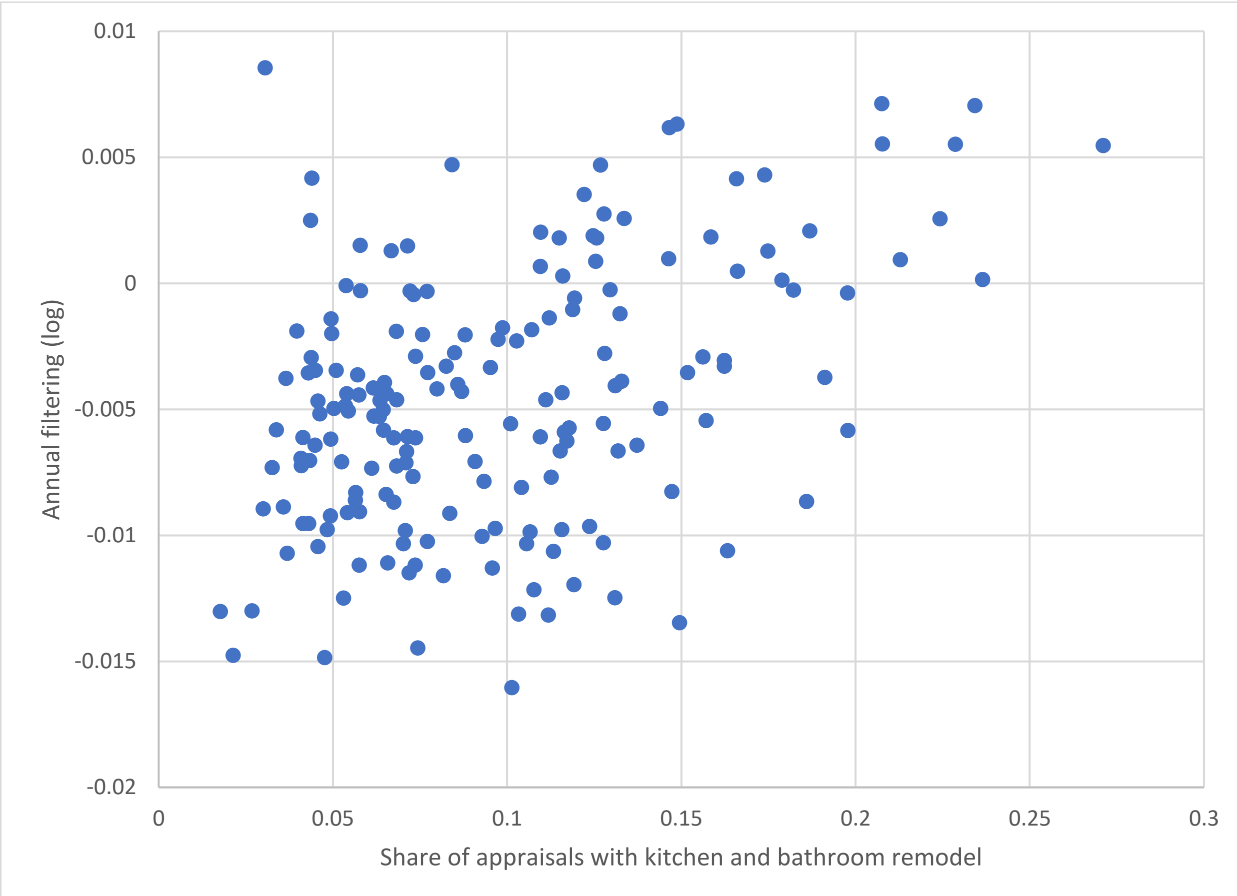
There does not appear to be much association between population growth and filtering rates.



Note: Log annual filtering rates are those reported in Table A1. Population estimates from Moody's Analytics based on US Census Bureau data.

e. By share of improvements

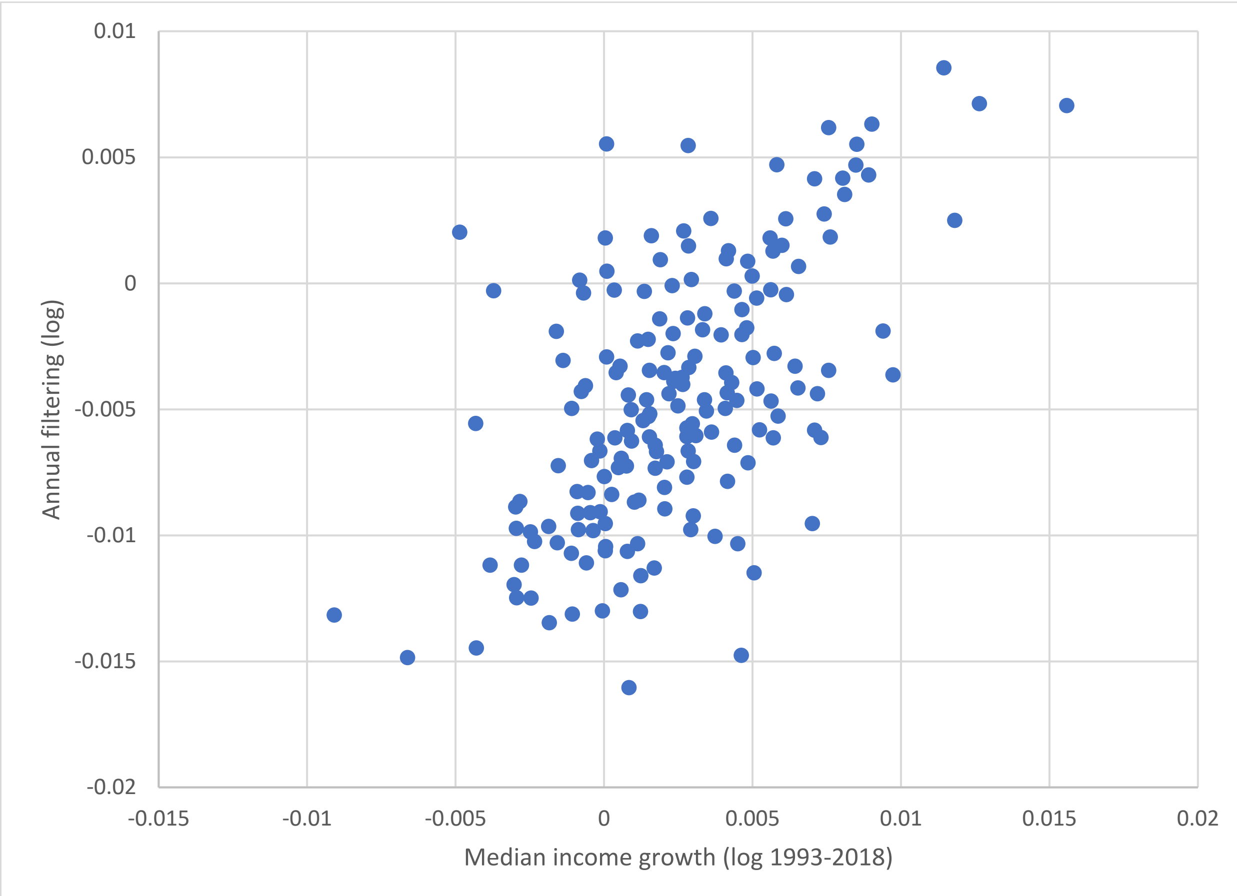
Filtering is positively correlated with improvements in appraisal data.



Note: Log annual filtering rates are those reported in Table A1. Appraisal data from the Uniform Mortgage Data Program. Remodel refers to significant finish and/or structural changes made within the past five years to both the kitchen and bathrooms.

f. By median real income growth

As expected, average filtering for an MSA is correlated with the rate of growth in median income over that period.



Note: Log annual filtering rates are those reported in Table A1. Median income data from Moody's Analytics based on U. S. Census Bureau data.

Table 3. Transitions of Reported Property Condition in Repeat Appraisals

a. Property condition transition rates in the top 30 downward filtering MSAs (percent)

Condition at First Appraisal	Condition at Second Appraisal			
	C1	C2	C3	C4
C1	7	54	37	1
C2	0	19	75	6
C3	0	4	79	17
C4	0	3	68	29

b. Property condition transition rates in the top 10 upwards filtering MSAs (percent)

Condition at First Appraisal	Condition at Second Appraisal			
	C1	C2	C3	C4
C1	7	60	32	1
C2	0	22	74	3
C3	0	9	82	10
C4	0	11	70	19

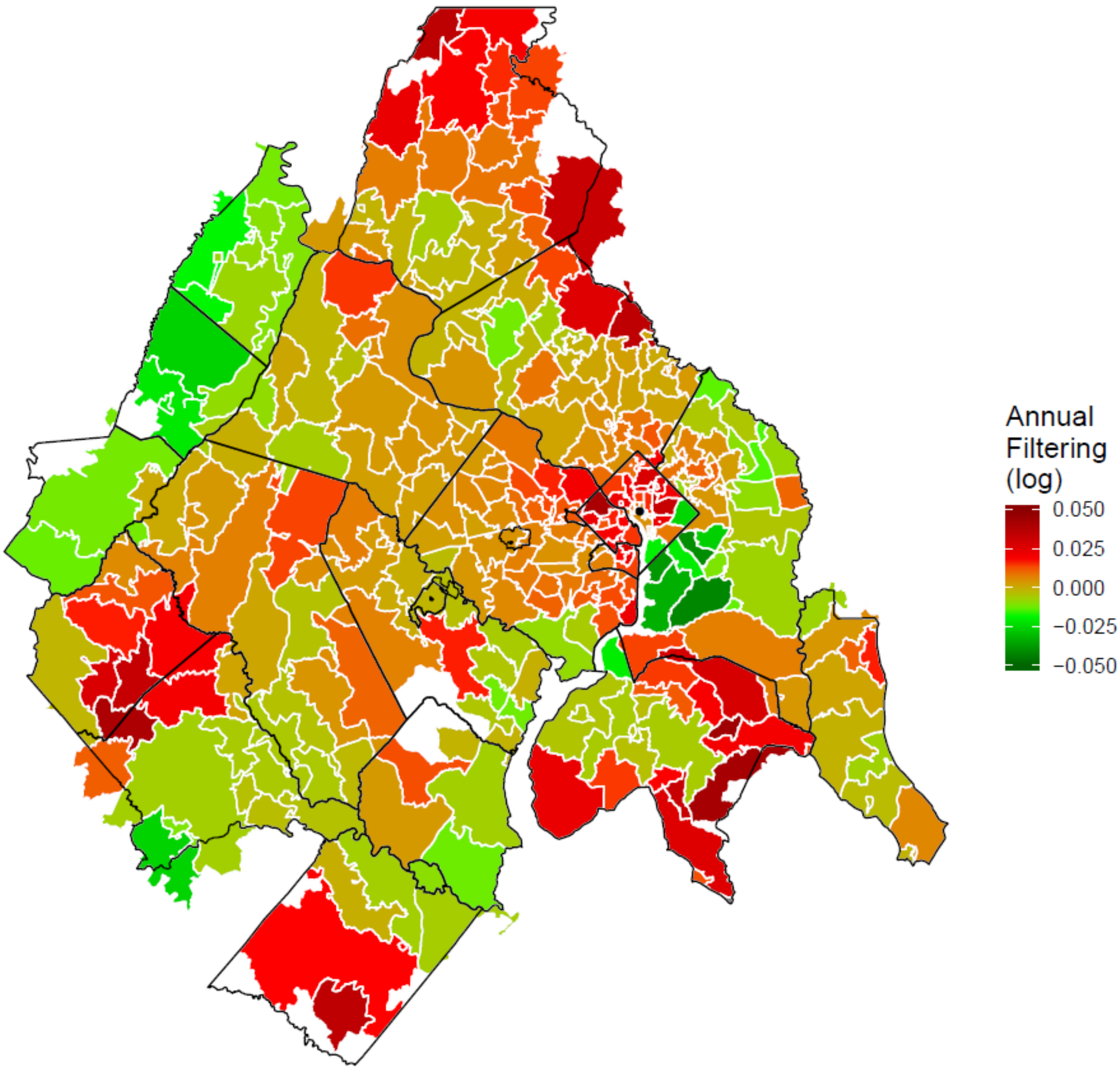
C1 = New construction/never occupied
C2 = Recently renovated/like new
C3 = Well maintained/limited depreciation
C4 = Adequately maintained/needs minimal repairs

Notes: The table reports the percentage of properties with a given condition at their second appraisal by the condition of the first appraisal. The table is based on pairs of purchase appraisals on the same property from the Uniform Appraisal Dataset for 1-unit single-family properties from 2012 through 2019 submitted for delivery to Freddie Mac. Top upward and downward filtering MSAs are based on the log annual filtering rates reported in Table A1.

Figure 4. Annual filtering (log) of ZIP Code centers using locally linear estimation

a. Washington, DC

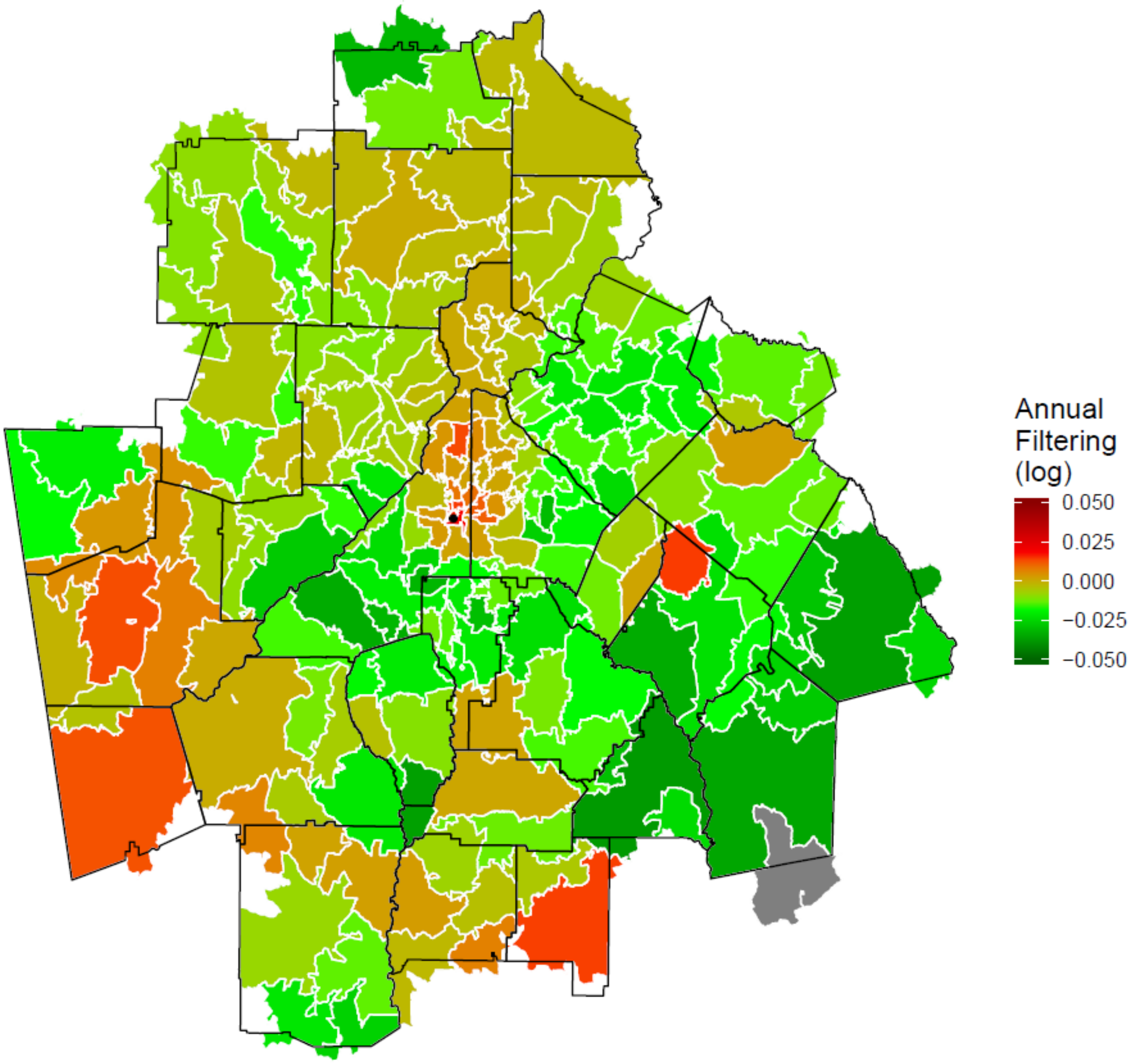
The Washington, DC metro area displays substantial heterogeneity in filtering rates across the MSA, with pockets of strong downward and upward filtering, while many ZIP Codes to the west have filtering rates close to zero.



Note: Each ZIP Code is colored based on the local linear estimate for the latitude and longitude of the centroid from the U. S. Census Bureau. The estimates use increasing bandwidths for areas farther from the city center that have fewer data. This helps preserve the heterogeneity for areas near the city center that would otherwise be washed out by the larger bandwidths. White borders denote ZIP Code boundaries. Black borders denote county boundaries. The small black dot denotes the coordinates of the central business district (CBD) from Holian and Kahn (2015).

b. Atlanta, Georgia

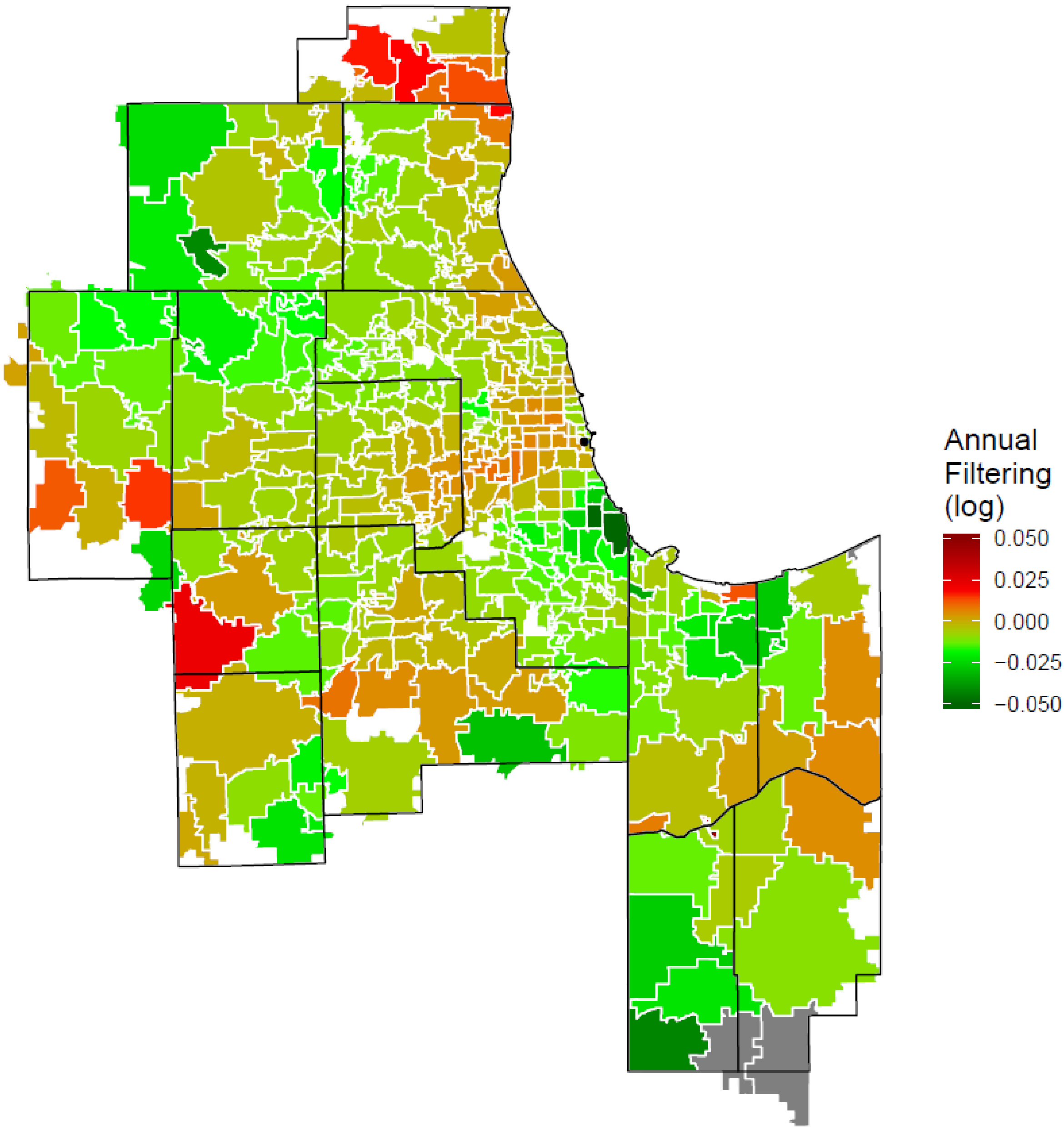
Atlanta has many regions of strong downward filtering, with upward filtering in the areas to the immediate north and east of the city center.



Note: Each ZIP Code is colored based on the local linear estimate for the latitude and longitude of the centroid from the U. S. Census Bureau. The estimates use increasing bandwidths for areas farther from the city center that have fewer data. This helps preserve the heterogeneity for areas near the city center that would otherwise be washed out by the larger bandwidths. White borders denote ZIP Code boundaries. Black borders denote county boundaries. The small black dot denotes the coordinates of the central business district (CBD) from Holian and Kahn (2015).

c. Chicago, Illinois

Chicago is dominated by regions of downward filtering, with some areas of upward filtering south west of the city center and on the fringes.



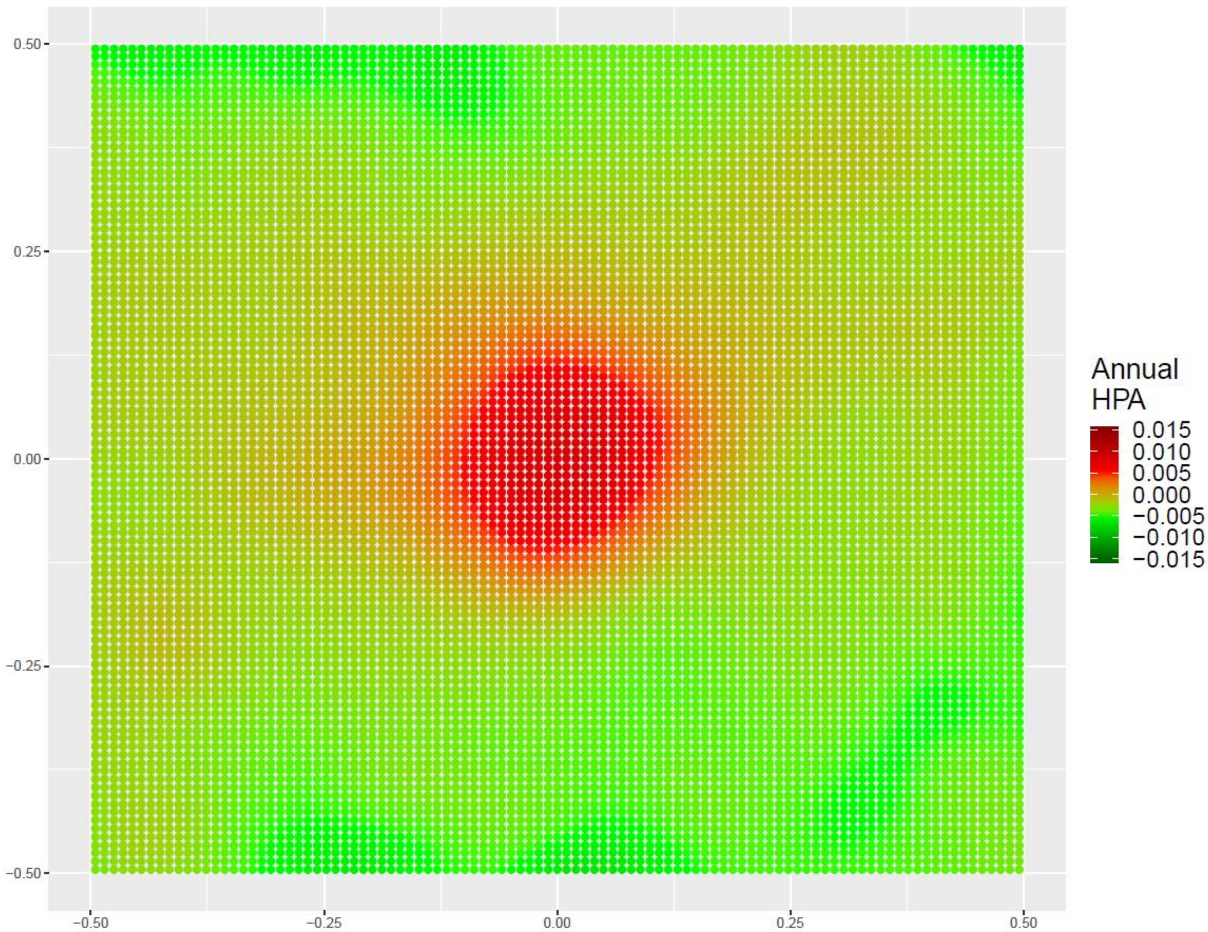
Note: Each ZIP Code is colored based on the local linear estimate for the latitude and longitude of the centroid from the U. S. Census Bureau. The estimates use increasing bandwidths for areas farther from the city center that have fewer data. This helps preserve the heterogeneity for areas near the city center that would otherwise be washed out by the larger bandwidths. White borders denote ZIP Code boundaries. Black borders denote county boundaries. The small black dot denotes the coordinates of the central business district (CBD) from Holian and Kahn (2015).

Figure 5. Within-city spatial variation

This figure shows the spatial deviation of ZIP Code house price appreciation (HPA) relative to the metropolitan statistical area (MSA), averaged over 26 MSAs.

a. House price appreciation

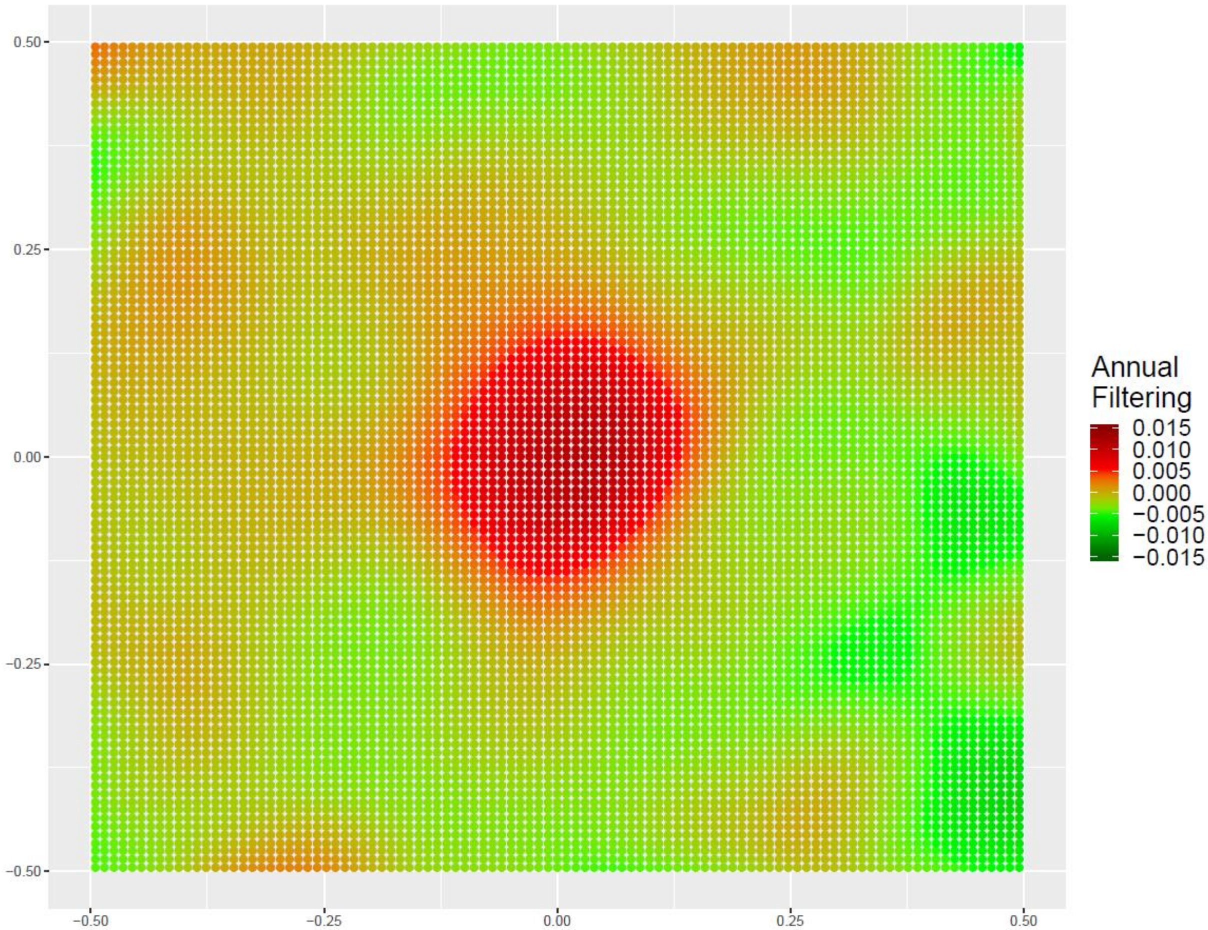
The area around the city center experiences the highest rates of house price growth, substantially above the average in the MSA, while the outer regions grow at about or just below the MSA average.



Note: The coordinate system is a normalized latitude and longitude where the location of the central business district (CBD) is (0,0) and the farthest ZIP Code centroid from the CBD in either direction defines the edge of the city at 1. The plot displays weighted local constant regression estimates with a bandwidth of 0.05 for a grid of point in the interior of the MSA half-way to the city's edge in either direction. We restrict the MSAs used to be consistent with those for which we have filtering estimates reported in Panel (b) of Figure 5, but this phenomenon is robust even for all MSAs for which we have indexes.

b. Annual filtering rates

The city centers have on average slower filtering rates than the MSA average, but there is substantial variation in filtering rates outside of the city center.



Note: Annual filtering rate is defined as log real income change divided by year between the transactions. The coordinate system is a normalized latitude and longitude where the location of the central business district (CBD) is (0,0) and the farthest ZIP Code centroid from the CBD in either direction defines the edge of the city at 1. The plot displays weighted local constant regression estimates using a bandwidth of 0.06 for a grid of point in the interior of the MSA half-way to the city's edge in either direction. The data include a combined sample of ZIP Code averages from 26 MSAs.

Table 4. Structural model of real change in income (log) by housing supply elasticity

	All	Elastic	Mid-elastic	Inelastic	Unclassified
First stage					
Percent change in HPI	0.611*** (0.0136)	0.616*** (0.0279)	0.644*** (0.0152)	0.605*** (0.0144)	0.569*** (0.0529)
KP weak instrument F-statistic	2024	488	1804	1762	116
Second stage					
Year between transactions	-0.0128*** (0.000505)	-0.0134*** (0.000593)	-0.0116*** (0.000908)	-0.0112*** (0.00116)	-0.0132*** (0.000296)
Change in log price	0.356*** (0.0132)	0.318*** (0.0248)	0.346*** (0.0171)	0.379*** (0.0282)	0.323*** (0.0215)
MSA fixed effects	380	155	79	18	128
Second-stage R-squared	0.045	0.034	0.05	0.068	0.036
Observations	1,233,887	482,027	352,526	218,264	181,070

Note: The table displays the structural model of real change in income (log) described in equation (7) stratified by supply elasticity. The percent change in the housing price index (HPI) is used as instrument for change in log house price. Standard errors clustered at the MSA level in parentheses. House supply elasticity estimates are from Saiz (2010). Inelastic metropolitan statistical areas (MSAs) are those with elasticity less than 1; mid-elastic refers to MSAs with elasticity between 1 and 2; and elastic MSAs are those with elasticity greater than 2. Unclassified refers to properties in MSAs without an estimated elasticity or not in an MSA. KP = Kleibergen-Paap weak instrument test. *** = Significant at 1 percent level. ** = Significant at 5 percent level. * = Significant at 10 percent level.

Table 5. Structural model of real change in income (log) by transaction year

	All	Pre-boom	Boom	Bust	Post-crisis
First stage					
Percent change in HPI	0.611*** (0.0136)	0.773*** (0.0221)	0.669*** (0.0159)	1.234*** (0.115)	0.575*** (0.0152)
KP weak instrument F-statistic	2024	1224	1764	115	1436
Second stage					
Year between transactions	-0.0128*** (0.000505)	-0.0133*** (0.00120)	-0.0166*** (0.00338)	-0.0111*** (0.00417)	0.000698 (0.00335)
Change in log price	0.356*** (0.0132)	0.384*** (0.0206)	0.358*** (0.0488)	0.652*** (0.0627)	0.435*** (0.0581)
MSA fixed effects	380	379	371	359	377
Second-stage R-squared	0.045	0.019	0.015	0.047	0.021
Observations	1,233,887	187,099	37,214	22,043	72,008

Note: The table displays the structural model of real change in income (log) described in equation (7) stratified by period. The percent change in the housing price index (HPI) is used as instrument for change in log house price. Standard errors clustered as the MSA level in parentheses. Data are partitioned, with only those pairs with both transactions in the same period included. The pre-boom period includes transactions from 1975 to December 2001; the boom period, from January 2002 to June 2006; the bust period, from July 2006 to December 2011; and the post-crisis period, from January 2012 to December 2018. KP = Kleibergen-Paap weak instrument test. *** = Significant at 1 percent level. ** = Significant at 5 percent level. * = Significant at 10 percent level.

Table 6. Structural model of real change in income (log) with time varying effect of years between transactions

First stage	
Percent change in HPI	0.541*** (0.0129)
KP weak instrument F-statistic	1770
Second stage	
Years between transactions in pre-boom period	-0.0172*** (0.000972)
Years between transactions in boom period	-0.0301*** (0.00108)
Years between transactions in bust period	-0.0102*** (0.000814)
Years between transactions in post-crisis period	-0.00632*** (0.000941)
Change in log price	0.455*** (0.0145)
MSA fixed effects	380
Second-stage R-squared	0.052
Observations	1,233,887

Note: This table presents an extension of the structural model of real change in income (log) described in equation (7) with differential depreciation rates across periods, but assumes the same relationship for house prices across periods. This specification allows us to use all the observation pairs in this analysis, unlike the results in Table 5. The percent change in house price index (HPI) is used as instrument for change in log house price. Standard errors clustered as the MSA level in parentheses. KP = Kleibergen-Paap weak instrument test. *** = Significant at 1 percent level. ** = Significant at 5 percent level. *= Significant at 10 percent level.

Table 7. Examining the impacts of gentrification on real change in income (log)

	Annual filtering rate	Log real income change	
Years between transactions	--	-0.00836*** (0.00134)	-0.0128*** (0.000502)
Change in log price	0.0479*** (0.00115)	0.435*** (0.0327)	0.357*** (0.0131)
In gentrification tract	0.00373*** (0.00130)	--	--
Gentrification tract × Years between transactions	--	--	0.00352*** (0.000860)
MSA fixed effects	380	215	380
Second-stage R-squared	0.013	0.07	0.045
First-stage coefficient on percent change in HPI	--	0.854*** (0.0228)	0.611*** (0.0136)
First-stage F statistics	--	1404	2011
Sample	Full Sample	Gentrification tracts	Full sample
Observations	1,233,888	12,520	1,233,887

Note: The first column displays the regression of the observation level filtering rates [see equation (5)] on the change in log prices and an indicator for being in a gentrification tract. The second column provides two-stage least squares (2SLS) estimates for the structural model restricted to only gentrification tracts. The third column provides 2SLS estimates for the structural model including an interaction term between "years between transactions" and being in a gentrification tract. Annual filtering rate is defined as log real income change divided by years between the transactions. Gentrification tract definition is from NCRC (2019). Standard errors clustered as the MSA level in parentheses. HPI = housing price index. *** = Significant at 1 percent level. ** = Significant at 5 percent level. * = Significant at 10 percent level.

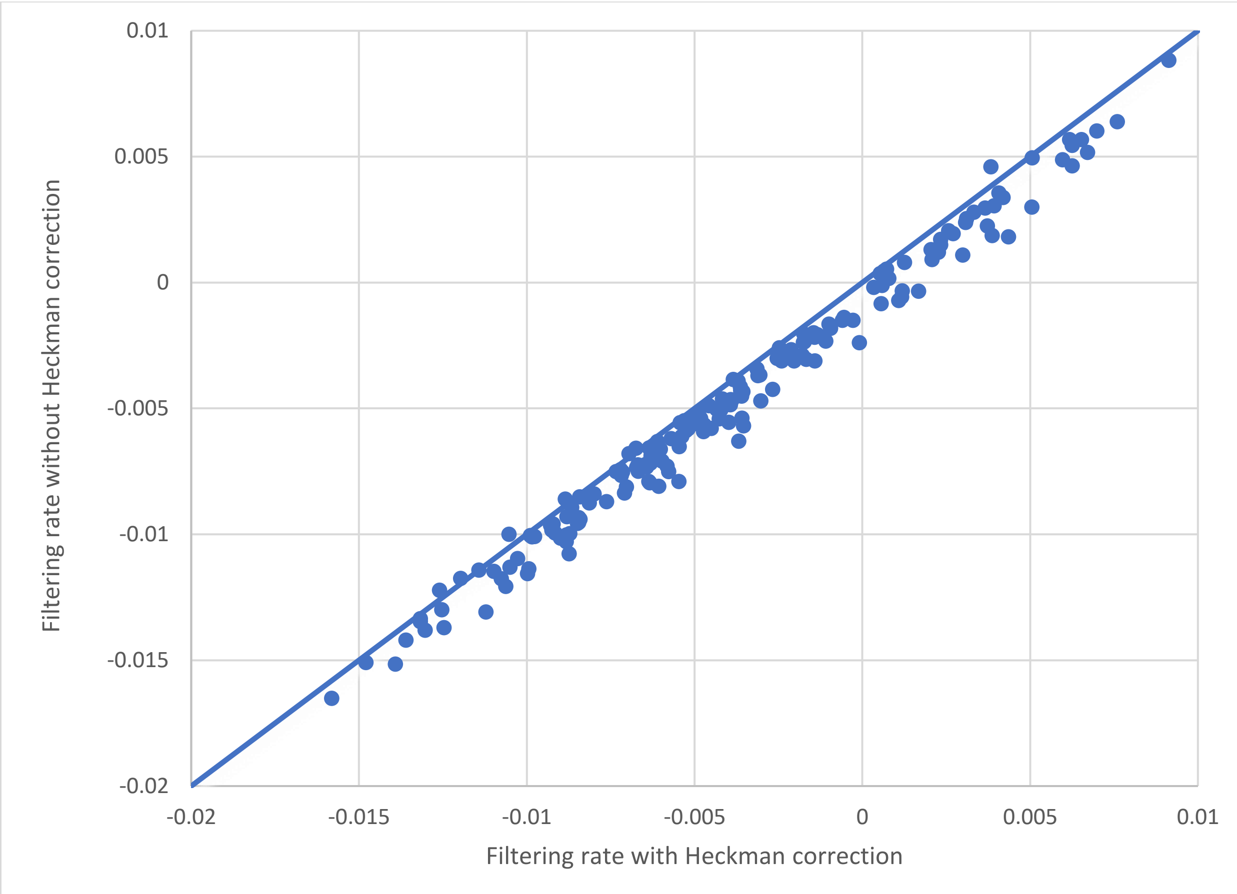
Table 8. Structural model of real change in income (log) with Heckman correction

	Without Heckman correction	With Heckman correction
First stage		
Percent change in HPI	0.630*** (0.0141)	0.629*** (0.0139)
KP weak instrument F-statistic	2003	2034
Second stage		
Years between transactions	-0.0124*** (0.000499)	-0.0116*** (0.000476)
Change in log price	0.353*** (0.0135)	0.339*** (0.0126)
Difference in inverse Mill's ratio	-- --	0.145*** (0.00956)
MSA fixed effects	380	380
Second-stage R-squared	0.044	0.046
Observations	1,051,519	1,051,519

Note: This table displays the structural model of real change in income (log) with Heckman correction described in equation (9) and without Heckman correction. The percent change in the housing price index (HPI) is used as instrument for change in log house price. Standard errors clustered as the MSA level in parentheses. KP = Kleibergen-Paap weak instrument test. *** = Significant at 1 percent level. ** = Significant at 5 percent level. * = Significant at 10 percent level.

Figure 6. Heckman correction: Linear model of real change in income (log) by MSA

The Heckman correction slightly decrease the filtering rate for most MSAs.



Note: The MSA filtering rate is estimated using equation (4), the linear model with no intercept. The horizontal axis shows the filtering rate with the Heckman correction term, $\beta(\lambda_{t+\tau,i} - \lambda_{t,i})$ for each MSA. The vertical axis shows the filtering rate without the Heckman correction term for the same MSA using the same sample.

Figure 7. Repeat income index estimates for construction age and effective age

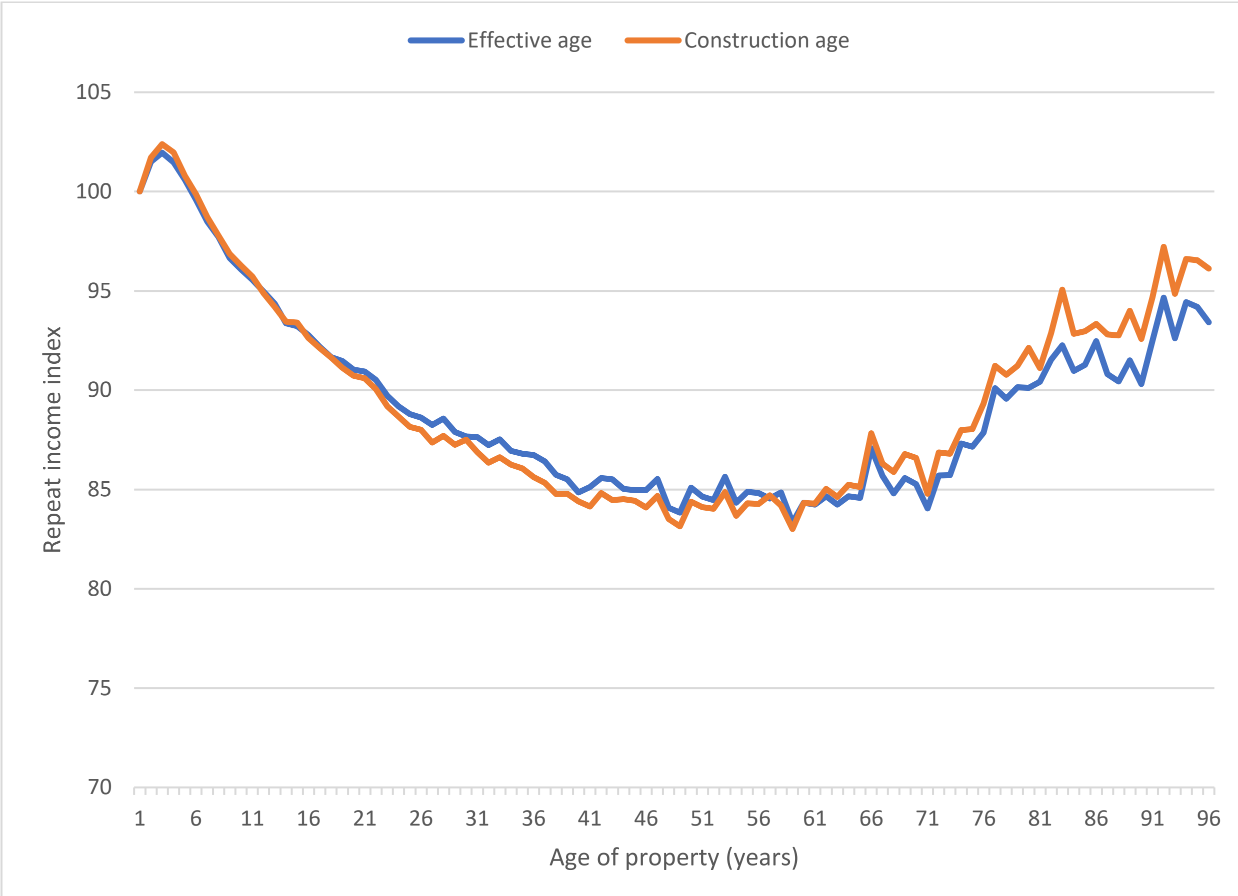


Figure 8. Comparison of the MSA filtering estimates for construction age and effective age

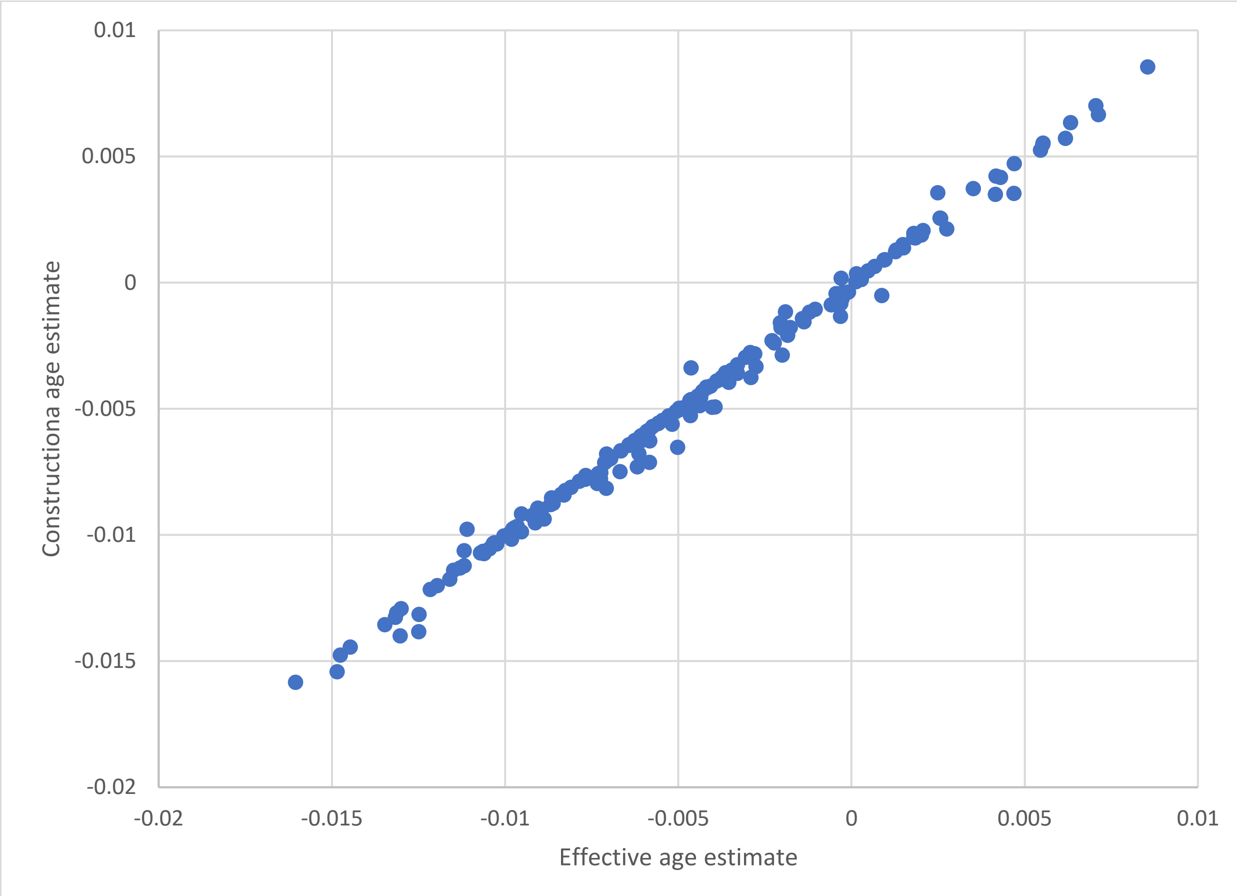
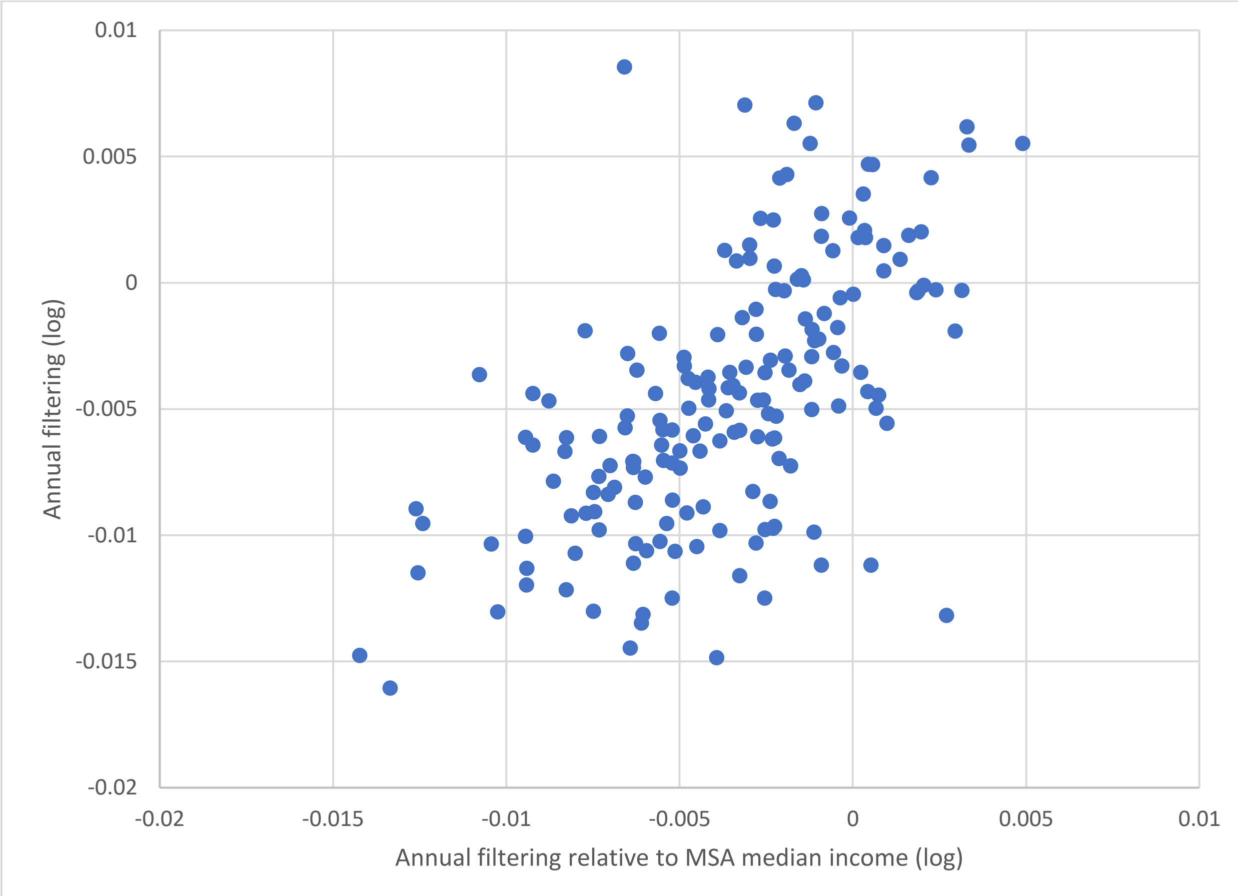


Figure 9. Scatterplot of the MSA filtering estimates with filtering relative to median income



Appendix

Table A1. Linear model of real change in income (log) by MSA

MSA	Filtering Rate (log)	Standard Error	40-Year (%)	MSA	Filtering Rate (log)	Standard Error	40-Year (%)	MSA	Filtering Rate (log)	Standard Error	40-Year (%)
Topeka, KS	-0.0161	0.0018	-47.38	Memphis, TN	-0.0070	0.0008	-24.27	Pittsburgh, PA	-0.0028	0.0007	-10.56
Macon, GA	-0.0149	0.0020	-44.79	Oklahoma City, OK	-0.0067	0.0008	-23.45	Salem, OR	-0.0028	0.0012	-10.45
Jackson, MS	-0.0148	0.0018	-44.59	Champaign, IL	-0.0067	0.0014	-23.39	Lancaster, PA	-0.0023	0.0010	-8.75
Fort Wayne, IN	-0.0145	0.0009	-43.94	Cincinnati, OH	-0.0067	0.0004	-23.36	York, PA	-0.0022	0.0012	-8.53
Toledo, OH	-0.0135	0.0008	-41.66	Baton Rouge, LA	-0.0064	0.0012	-22.65	Ogden, UT	-0.0021	0.0009	-7.87
Flint, MI	-0.0132	0.0015	-40.95	Davenport, IA	-0.0064	0.0010	-22.65	Madison, WI	-0.0020	0.0008	-7.84
South Bend, IN	-0.0131	0.0013	-40.86	Reading, PA	-0.0063	0.0014	-22.15	Gainesville, FL	-0.0020	0.0019	-7.69
Myrtle Beach, SC	-0.0130	0.0021	-40.62	Boise City, ID	-0.0062	0.0011	-21.90	Port St. Lucie, FL	-0.0019	0.0019	-7.36
Spartanburg, SC	-0.0130	0.0018	-40.55	Orlando, FL	-0.0061	0.0006	-21.78	Fargo, ND	-0.0019	0.0015	-7.32
Greensboro, NC	-0.0125	0.0008	-39.32	Des Moines, IA	-0.0061	0.0007	-21.75	Richmond, VA	-0.0018	0.0007	-7.10
Youngstown, OH	-0.0125	0.0017	-39.30	Fayetteville, AR	-0.0061	0.0016	-21.71	Phoenix, AZ	-0.0018	0.0004	-6.84
Evansville, IN	-0.0122	0.0009	-38.52	Ann Arbor, MI	-0.0061	0.0012	-21.62	Fresno, CA	-0.0014	0.0013	-5.52
Terre Haute, IN	-0.0120	0.0018	-38.02	Stockton, CA	-0.0061	0.0012	-21.59	Spokane, WA	-0.0014	0.0011	-5.33
Elkhart, IN	-0.0116	0.0017	-37.12	Lincoln, NE	-0.0061	0.0012	-21.49	Tampa, FL	-0.0012	0.0005	-4.72
Lawrence, KS	-0.0115	0.0020	-36.85	Louisville, KY	-0.0059	0.0006	-21.05	Worcester, MA	-0.0011	0.0008	-4.11
Owensboro, KY	-0.0113	0.0016	-36.36	Chicago, IL	-0.0058	0.0002	-20.83	Minneapolis, MN	-0.0006	0.0003	-2.32
Hickory, NC	-0.0112	0.0016	-36.06	St. Cloud, MN	-0.0058	0.0013	-20.80	Louisville, KY	-0.0004	0.0011	-1.76
Winston, NC	-0.0112	0.0010	-36.06	Raleigh, NC	-0.0058	0.0005	-20.77	New Haven, CT	-0.0004	0.0009	-1.53
Mobile, AL	-0.0111	0.0018	-35.85	Riverside, CA	-0.0057	0.0006	-20.51	Asheville, NC	-0.0003	0.0016	-1.28
Kennewick, WA	-0.0107	0.0014	-34.84	Tucson, AZ	-0.0056	0.0009	-20.00	Olympia, WA	-0.0003	0.0017	-1.24
Grand Rapids, MI	-0.0106	0.0006	-34.66	Rochester, NY	-0.0056	0.0007	-19.94	Palm Bay, FL	-0.0003	0.0012	-1.18
Akron, OH	-0.0106	0.0009	-34.58	Waterloo, IA	-0.0055	0.0018	-19.59	Miami, FL	-0.0003	0.0004	-1.09
Las Vegas, NV	-0.0105	0.0007	-34.16	Lynchburg, VA	-0.0053	0.0021	-19.04	Manchester, NH	-0.0003	0.0010	-1.01
Eau Claire, WI	-0.0103	0.0018	-33.87	Omaha, NE	-0.0053	0.0008	-19.01	Eugene, OR	-0.0001	0.0011	-0.36
Green Bay, WI	-0.0103	0.0015	-33.85	Tallahassee, FL	-0.0052	0.0013	-18.71	Urban Honolulu, HI	0.0001	0.0015	0.51
Kalamazoo, MI	-0.0103	0.0012	-33.77	Durham, NC	-0.0051	0.0009	-18.36	Bridgeport, CT	0.0001	0.0009	0.58
Wichita, KS	-0.0102	0.0011	-33.61	Cape Coral, FL	-0.0050	0.0014	-18.19	Sacramento, CA	0.0003	0.0006	1.14
Bloomington, IL	-0.0100	0.0012	-33.08	Modesto, CA	-0.0050	0.0013	-18.03	Trenton, NJ	0.0005	0.0014	1.92
Detroit, MI	-0.0099	0.0003	-32.62	Milwaukee, WI	-0.0050	0.0006	-18.03	Portland, ME	0.0007	0.0013	2.70
Indianapolis, IN	-0.0098	0.0005	-32.46	Charlotte, NC	-0.0049	0.0005	-17.70	Billings, MT	0.0009	0.0019	3.54
Springfield, MO	-0.0098	0.0012	-32.38	Lafayette, LA	-0.0047	0.0017	-17.04	Philadelphia, PA	0.0009	0.0003	3.82
Rockford, IL	-0.0098	0.0011	-32.35	Nashville, TN	-0.0047	0.0006	-16.97	Albany, NY	0.0010	0.0010	3.95
Lansing, MI	-0.0097	0.0010	-32.21	Athens, GA	-0.0046	0.0020	-16.91	Providence, RI	0.0013	0.0008	5.21
Racine, WI	-0.0097	0.0015	-32.02	Allentown, PA	-0.0046	0.0009	-16.91	State College, PA	0.0013	0.0019	5.30
Columbia, SC	-0.0095	0.0009	-31.70	Reno, NV	-0.0044	0.0012	-16.27	Bellingham, WA	0.0015	0.0015	6.10
Tuscaloosa, AL	-0.0095	0.0020	-31.70	Provo, UT	-0.0044	0.0011	-16.07	Charleston, SC	0.0015	0.0010	6.18
Bowling Green, KY	-0.0092	0.0018	-30.87	Pensacola, FL	-0.0044	0.0016	-16.07	Burlington, VT	0.0018	0.0013	7.47
Columbus, GA	-0.0091	0.0019	-30.59	Columbus, OH	-0.0044	0.0005	-15.97	Duluth, MN	0.0018	0.0014	7.47
Greenville, SC	-0.0091	0.0007	-30.54	Albuquerque, NM	-0.0043	0.0009	-15.80	Denver, CO	0.0018	0.0004	7.64
Beaumont, TX	-0.0091	0.0017	-30.40	Colorado Springs, CO	-0.0042	0.0010	-15.43	Vallejo, CA	0.0019	0.0015	7.81
Lubbock, TX	-0.0090	0.0017	-30.09	Mankato, MN	-0.0042	0.0021	-15.30	Anchorage, AK	0.0020	0.0013	8.42
Augusta, GA	-0.0089	0.0013	-29.87	Syracuse, NY	-0.0041	0.0013	-15.02	New York, NY	0.0021	0.0003	8.63
Chattanooga, TN	-0.0087	0.0012	-29.36	Jacksonville, FL	-0.0040	0.0009	-14.85	Austin, TX	0.0025	0.0005	10.47
Cleveland, OH	-0.0087	0.0007	-29.28	Savannah, GA	-0.0039	0.0013	-14.58	Baltimore, MD	0.0026	0.0005	10.78
Little Rock, AR	-0.0086	0.0011	-29.14	St. Louis, MO	-0.0039	0.0004	-14.41	Virginia Beach, VA	0.0026	0.0009	10.83
Columbia, MO	-0.0084	0.0014	-28.48	Houston, TX	-0.0038	0.0003	-14.03	Salt Lake City, UT	0.0028	0.0011	11.63
Tulsa, OK	-0.0083	0.0009	-28.25	New Orleans, LA	-0.0037	0.0011	-13.89	Portland, OR	0.0035	0.0004	15.12
Canton, OH	-0.0083	0.0015	-28.14	Greeley, CO	-0.0036	0.0014	-13.52	Washington, DC	0.0042	0.0003	18.06
Peoria, IL	-0.0081	0.0012	-27.67	Dallas, TX	-0.0036	0.0003	-13.27	Bend, OR	0.0042	0.0019	18.15
Cedar Rapids, IA	-0.0079	0.0011	-26.98	Deltona, FL	-0.0035	0.0015	-13.20	Boston, MA	0.0043	0.0004	18.77
Kansas City, MO	-0.0077	0.0004	-26.48	Buffalo, NY	-0.0035	0.0009	-13.20	Boulder, CO	0.0047	0.0009	20.64
Huntington, WV	-0.0077	0.0019	-26.42	Wilmington, NC	-0.0035	0.0013	-12.92	Charlottesville, VA	0.0047	0.0018	20.68
Lakeland, FL	-0.0073	0.0015	-25.44	Sioux Falls, SD	-0.0035	0.0017	-12.92	Los Angeles, CA	0.0055	0.0004	24.41
El Paso, TX	-0.0073	0.0019	-25.35	Harrisburg, PA	-0.0033	0.0011	-12.51	San Diego, CA	0.0055	0.0006	24.71
Atlanta, GA	-0.0073	0.0003	-25.17	Rochester, MN	-0.0033	0.0011	-12.33	Oxnard, CA	0.0055	0.0010	24.76
Bakersfield, CA	-0.0072	0.0014	-25.14	Roanoke, VA	-0.0033	0.0013	-12.33	Santa Rosa, CA	0.0062	0.0013	28.04
Lexington, KY	-0.0071	0.0008	-24.81	Hartford, CT	-0.0031	0.0007	-11.52	Seattle, WA	0.0063	0.0004	28.76
Birmingham, AL	-0.0071	0.0008	-24.66	San Antonio, TX	-0.0030	0.0007	-11.13	San Jose, CA	0.0071	0.0008	32.58
Knoxville, TN	-0.0071	0.0009	-24.63	Springfield, MA	-0.0029	0.0015	-11.02	San Francisco, CA	0.0071	0.0006	33.00
Huntsville, AL	-0.0070	0.0012	-24.51	North Port, FL	-0.0029	0.0013	-10.95	Midland, TX	0.0086	0.0015	40.78

Note: The filtering rate is estimated using equation (4), the linear model with no intercept. The 40-year percentage change is computed as $[\exp(40 * Filtering\ coefficient) - 1] * 100$. Only the first city of each MSA is listed.

Table A2. Heckman selection first stage summary statistics

Variable	Mean
GSE indicator	0.41
FICO	740
LTV	79.89
Loan amount	170115
DTI	35.31
One borrower	0.50
One unit	0.98
Fixed product	0.63
30-yr product	0.80
First time home buyer	0.50
Sample size	4,465,886

Year	Sample size	Year	Sample size
1995	131,019	2007	193,532
1996	169,823	2008	144,100
1997	195,329	2009	133,607
1998	253,600	2010	114,156
1999	251,110	2011	106,478
2000	227,814	2012	117,086
2001	242,300	2013	133,441
2002	240,217	2014	133,059
2003	251,286	2015	152,466
2004	242,838	2016	171,802
2005	256,140	2017	180,278
2006	241,783	2018	182,622

Table A3. Heckman selection first stage probit regression

Dependent Variable: Indicator of GSE loan					
Independent Variable	Coefficient	Standard Error	Independent Variable	Coefficient	Standard Error
FICO [650,700)	0.082	(0.004)	Year1996	-0.058	(0.005)
FICO [700,750)	0.330	(0.004)	Year1997	-0.225	(0.005)
FICO [750,800)	0.618	(0.004)	Year1998	-0.173	(0.005)
FICO >=800	0.746	(0.005)	Year1999	-0.206	(0.005)
LTV [80,95]	0.039	(0.002)	Year2000	-0.235	(0.005)
LTV (95,105)	-0.749	(0.002)	Year2001	-0.044	(0.005)
Loan amount (5k,10k]	0.782	(0.003)	Year2002	-0.039	(0.005)
Loan amount (10k,20k]	1.004	(0.003)	Year2003	0.021	(0.005)
Loan amount (20k,40k]	0.977	(0.003)	Year2004	-0.076	(0.005)
Loan amount (40k,750k]	0.606	(0.017)	Year2005	-0.176	(0.005)
Loan amount >750k	-1.730	(0.031)	Year2006	-0.103	(0.005)
DTI [30,43)	-0.019	(0.002)	Year2007	0.160	(0.005)
DTI [43,50)	-0.044	(0.002)	Year2008	-0.056	(0.006)
DTI [50,70]	-0.034	(0.003)	Year2009	-0.149	(0.006)
One borrower indicator	0.078	(0.002)	Year2010	-0.108	(0.006)
One unit indicator	0.327	(0.006)	Year2011	-0.066	(0.006)
Fixed product	1.517	(0.002)	Year2012	-0.050	(0.006)
30 Yr product	0.830	(0.002)	Year2013	-0.009	(0.006)
First time home buyer	0.086	(0.002)	Year2014	-0.019	(0.006)
UPB>40k × one unit	-0.397	(0.017)	Year2015	-0.068	(0.006)
Constant	-3.326	(0.009)	Year2016	-0.030	(0.006)
Sample size	4,465,886		Year2017	0.001	(0.006)
R square	0.348		Year2018	-0.164	(0.006)

Notes: This table provides probit regression estimates for the selection model described in equation (8) using NMDB data. The base categories in this regression are for FICO, FICO<650; for UPB, OUPB<10k; for DTI DTI<30%; for LTV, LTV<80; and for origination year the base is 1995. All estimates are significant at 1 percent except year 2013 and year 2017 indicator.
We had also tried separate probit regression for each origination year and obtain very similar results for the second stage filtering rate estimates.

Table A4. Linear model of squared residuals on time between sales by MSA

MSA	Intercept	Linear Coeff	8-Year Var	MSA	Intercept	Linear Coeff	8-Year Var	MSA	Intercept	Linear Coeff	8-Year Var
Akron, OH	0.2292	0.0066	0.28	Fort Wayne, IN	0.2406	0.0063	0.29	Peoria, IL	0.2257	0.0068	0.28
Albany, NY	0.2358	0.0040	0.27	Fresno, CA	0.3015	0.0081	0.37	Philadelphia, PA	0.2378	0.0051	0.28
Albuquerque, NM	0.2641	0.0090	0.34	Gainesville, FL	0.3834	0.0005	0.39	Phoenix, AZ	0.2968	0.0062	0.35
Allentown, PA	0.2070	0.0074	0.27	Grand Rapids, MI	0.2513	0.0048	0.29	Pittsburgh, PA	0.2572	0.0047	0.29
Anchorage, AK	0.2286	0.0051	0.27	Greeley, CO	0.2527	0.0059	0.30	Port St. Lucie, FL	0.3755	0.0014	0.39
Ann Arbor, MI	0.2317	0.0078	0.29	Green Bay, WI	0.2552	0.0011	0.26	Portland, ME	0.2956	0.0003	0.30
Asheville, NC	0.3100	0.0102	0.39	Greensboro, NC	0.2614	0.0052	0.30	Portland, OR	0.2577	0.0060	0.31
Athens, GA	0.2491	0.0160	0.38	Greenville, SC	0.2403	0.0066	0.29	Providence, RI	0.2258	0.0032	0.25
Atlanta, GA	0.2287	0.0085	0.30	Harrisburg, PA	0.2365	0.0046	0.27	Provo, UT	0.2173	0.0113	0.31
Augusta, GA	0.2740	0.0074	0.33	Hartford, CT	0.2229	0.0047	0.26	Racine, WI	0.2556	-0.0008	0.25
Austin, TX	0.2334	0.0063	0.28	Hickory, NC	0.2339	0.0088	0.30	Raleigh, NC	0.2449	0.0046	0.28
Bakersfield, CA	0.2629	0.0073	0.32	Houston, TX	0.2269	0.0062	0.28	Reading, PA	0.2456	-0.0002	0.24
Baltimore, MD	0.2526	0.0036	0.28	Huntington, WV	0.2618	0.0050	0.30	Reno, NV	0.3049	0.0055	0.35
Baton Rouge, LA	0.2555	0.0054	0.30	Huntsville, AL	0.2391	0.0058	0.29	Richmond, VA	0.2423	0.0064	0.29
Beaumont, TX	0.2569	0.0034	0.28	Indianapolis, IN	0.2465	0.0056	0.29	Riverside, CA	0.3293	0.0018	0.34
Bellingham, WA	0.2854	0.0062	0.33	Jackson, MS	0.2298	0.0082	0.30	Roanoke, VA	0.2594	0.0062	0.31
Bend, OR	0.3596	0.0084	0.43	Jacksonville, FL	0.2581	0.0083	0.32	Rochester, MN	0.2821	0.0028	0.30
Billings, MT	0.2489	0.0084	0.32	Kalamazoo, MI	0.2029	0.0086	0.27	Rochester, NY	0.2477	0.0039	0.28
Birmingham, AL	0.2661	0.0067	0.32	Kansas City, MO	0.2223	0.0117	0.32	Rockford, IL	0.2103	0.0069	0.27
Bloomington, IL	0.2272	0.0027	0.25	Kennewick, WA	0.2289	0.0070	0.29	Sacramento, CA	0.3133	0.0039	0.34
Boise City, ID	0.3423	0.0088	0.41	Knoxville, TN	0.2407	0.0076	0.30	Salem, OR	0.2754	0.0054	0.32
Boston, MA	0.2331	0.0025	0.25	Lafayette, LA	0.2278	0.0107	0.31	Salt Lake City, UT	0.2754	0.0040	0.31
Boulder, CO	0.2720	0.0053	0.31	Lakeland, FL	0.2674	0.0065	0.32	San Antonio, TX	0.2576	0.0053	0.30
Bowling Green, KY	0.2367	0.0071	0.29	Lancaster, PA	0.2695	0.0014	0.28	San Diego, CA	0.2899	0.0019	0.30
Bridgeport, CT	0.2287	0.0050	0.27	Lansing, MI	0.2800	0.0014	0.29	San Francisco, CA	0.2715	0.0035	0.30
Buffalo, NY	0.2570	0.0028	0.28	Las Vegas, NV	0.3318	0.0054	0.38	San Jose, CA	0.2435	0.0038	0.27
Burlington, VT	0.2697	0.0032	0.30	Lawrence, KS	0.2500	0.0173	0.39	Santa Rosa, CA	0.2922	0.0067	0.35
Canton, OH	0.2499	0.0038	0.28	Lexington, KY	0.2442	0.0073	0.30	Savannah, GA	0.2254	0.0118	0.32
Cape Coral, FL	0.3265	0.0074	0.39	Lincoln, NE	0.2602	0.0012	0.27	Seattle, WA	0.2500	0.0047	0.29
Cedar Rapids, IA	0.1982	0.0044	0.23	Little Rock, AR	0.2786	0.0061	0.33	Sioux Falls, SD	0.2495	0.0041	0.28
Champaign, IL	0.3214	0.0026	0.34	Los Angeles, CA	0.2649	0.0049	0.30	South Bend, IN	0.2226	0.0088	0.29
Charleston, SC	0.2825	0.0055	0.33	Louisville, KY	0.2489	0.0074	0.31	Spartanburg, SC	0.2691	0.0071	0.33
Charlotte, NC	0.2494	0.0069	0.30	Lubbock, TX	0.2630	0.0076	0.32	Spokane, WA	0.2989	0.0059	0.35
Charlottesville, VA	0.2843	0.0051	0.33	Lynchburg, VA	0.2752	0.0068	0.33	Springfield, MA	0.2418	0.0031	0.27
Chattanooga, TN	0.2438	0.0062	0.29	Macon, GA	0.2328	0.0092	0.31	Springfield, MO	0.3125	0.0036	0.34
Chicago, IL	0.2297	0.0041	0.26	Madison, WI	0.2417	0.0033	0.27	St. Cloud, MN	0.2968	0.0011	0.31
Cincinnati, OH	0.2400	0.0055	0.28	Manchester, NH	0.2447	0.0006	0.25	St. Louis, MO	0.2456	0.0053	0.29
Cleveland, OH	0.2340	0.0063	0.28	Mankato, MN	0.2397	0.0069	0.29	State College, PA	0.2041	0.0076	0.27
Colorado Springs, CO	0.2617	0.0048	0.30	Memphis, TN	0.2220	0.0083	0.29	Stockton, CA	0.2909	0.0034	0.32
Columbia, MO	0.2797	0.0055	0.32	Miami, FL	0.2666	0.0052	0.31	Syracuse, NY	0.2482	0.0017	0.26
Columbia, SC	0.2549	0.0086	0.32	Midland, TX	0.2437	0.0020	0.26	Tallahassee, FL	0.2593	0.0085	0.33
Columbus, GA	0.2339	0.0081	0.30	Milwaukee, WI	0.2387	0.0026	0.26	Tampa, FL	0.2761	0.0063	0.33
Columbus, OH	0.2451	0.0041	0.28	Minneapolis, MN	0.2411	0.0051	0.28	Terre Haute, IN	0.2792	0.0028	0.30
Dallas, TX	0.2339	0.0054	0.28	Mobile, AL	0.2897	0.0031	0.31	Toledo, OH	0.2663	0.0038	0.30
Davenport, IA	0.2068	0.0091	0.28	Modesto, CA	0.3317	0.0003	0.33	Topeka, KS	0.1551	0.0282	0.38
Deltona, FL	0.3221	0.0070	0.38	Myrtle Beach, SC	0.3351	0.0096	0.41	Trenton, NJ	0.2272	0.0034	0.25
Denver, CO	0.2646	0.0050	0.30	Nashville, TN	0.2591	0.0067	0.31	Tucson, AZ	0.3624	0.0025	0.38
Des Moines, IA	0.2132	0.0032	0.24	New Haven, CT	0.2184	0.0039	0.25	Tulsa, OK	0.2570	0.0072	0.31
Detroit, MI	0.2293	0.0047	0.27	New Orleans, LA	0.2520	0.0069	0.31	Tuscaloosa, AL	0.3055	0.0002	0.31
Duluth, MN	0.3047	0.0043	0.34	New York, NY	0.2339	0.0028	0.26	Urban Honolulu, HI	0.2913	-0.0009	0.28
Durham, NC	0.2871	0.0055	0.33	North Port, FL	0.3207	0.0068	0.37	Vallejo, CA	0.2963	0.0024	0.32
Eau Claire, WI	0.1934	0.0105	0.28	Ogden, UT	0.2388	0.0045	0.27	Virginia Beach, VA	0.2170	0.0080	0.28
El Paso, TX	0.2407	0.0112	0.33	Oklahoma City, OK	0.2634	0.0069	0.32	Washington, DC	0.2316	0.0049	0.27
Elkhart, IN	0.3200	-0.0017	0.31	Olympia, WA	0.2294	0.0083	0.30	Waterloo, IA	0.1843	0.0106	0.27
Eugene, OR	0.2711	0.0080	0.33	Omaha, NE	0.2384	0.0032	0.26	Wichita, KS	0.2191	0.0095	0.30
Evansville, IN	0.2511	0.0056	0.30	Orlando, FL	0.2825	0.0053	0.32	Wilmington, NC	0.3191	0.0047	0.36
Fargo, ND	0.2204	0.0058	0.27	Owensboro, KY	0.2897	0.0050	0.33	Winston, NC	0.2673	0.0041	0.30
Fayetteville, AR	0.2221	0.0104	0.31	Oxnard, CA	0.2640	0.0030	0.29	Worcester, MA	0.2205	0.0037	0.25
Flint, MI	0.2520	0.0063	0.30	Palm Bay, FL	0.2693	0.0088	0.34	York, PA	0.2217	0.0064	0.27
Fort Collins, CO	0.2839	0.0059	0.33	Pensacola, FL	0.2656	0.0104	0.35	Youngstown, OH	0.2478	0.0057	0.29

Note: The residuals used in this estimation are from the model using equation (4) and whose coefficients are reported in Table A1. The 8-year variance is computed by applying the linear model for that MSA to a time between sales of eight years, the national mean. Only the first city of each MSA is listed.

Table A5. Comparison of the MSA filtering estimates with filtering relative to median income

MSA	Filtering Rate (log)	Relative to Median (log)	MSA	Filtering Rate (log)	Relative to Median (log)	MSA	Filtering Rate (log)	Relative to Median (log)
Topeka, KS	-0.0161	-0.0134	Memphis, TN	-0.0070	-0.0021	Pittsburgh, PA	-0.0028	-0.0065
Macon, GA	-0.0149	-0.0039	Oklahoma City, OK	-0.0067	-0.0083	Salem, OR	-0.0028	-0.0006
Jackson, MS	-0.0148	-0.0142	Champaign, IL	-0.0067	-0.0044	Lancaster, PA	-0.0023	-0.0011
Fort Wayne, IN	-0.0145	-0.0064	Cincinnati, OH	-0.0067	-0.0050	York, PA	-0.0022	-0.0010
Toledo, OH	-0.0135	-0.0061	Baton Rouge, LA	-0.0064	-0.0092	Ogden, UT	-0.0021	-0.0039
Flint, MI	-0.0132	0.0027	Davenport, IA	-0.0064	-0.0055	Madison, WI	-0.0020	-0.0028
South Bend, IN	-0.0131	-0.0061	Reading, PA	-0.0063	-0.0038	Gainesville, FL	-0.0020	-0.0056
Myrtle Beach, SC	-0.0130	-0.0103	Boise City, ID	-0.0062	-0.0023	Port St. Lucie, FL	-0.0019	0.0030
Spartanburg, SC	-0.0130	-0.0075	Orlando, FL	-0.0061	-0.0023	Fargo, ND	-0.0019	-0.0077
Greensboro, NC	-0.0125	-0.0025	Des Moines, IA	-0.0061	-0.0083	Richmond, VA	-0.0018	-0.0012
Youngstown, OH	-0.0125	-0.0052	Fayetteville, AR	-0.0061	-0.0094	Phoenix, AZ	-0.0018	-0.0004
Evansville, IN	-0.0122	-0.0083	Ann Arbor, MI	-0.0061	-0.0028	Fresno, CA	-0.0014	-0.0014
Terre Haute, IN	-0.0120	-0.0094	Stockton, CA	-0.0061	-0.0073	Spokane, WA	-0.0014	-0.0032
Elkhart, IN	-0.0116	-0.0033	Lincoln, NE	-0.0061	-0.0046	Tampa, FL	-0.0012	-0.0008
Lawrence, KS	-0.0115	-0.0126	Louisville, KY	-0.0059	-0.0034	Worcester, MA	-0.0011	-0.0028
Owensboro, KY	-0.0113	-0.0094	Chicago, IL	-0.0058	-0.0033	Minneapolis, MN	-0.0006	-0.0004
Hickory, NC	-0.0112	0.0005	St. Cloud, MN	-0.0058	-0.0052	Fort Collins, CO	-0.0004	0.0000
Winston, NC	-0.0112	-0.0009	Raleigh, NC	-0.0058	-0.0055	New Haven, CT	-0.0004	0.0018
Mobile, AL	-0.0111	-0.0063	Riverside, CA	-0.0057	-0.0066	Asheville, NC	-0.0003	0.0019
Kennewick, WA	-0.0107	-0.0080	Tucson, AZ	-0.0056	-0.0043	Olympia, WA	-0.0003	-0.0020
Grand Rapids, MI	-0.0106	-0.0051	Rochester, NY	-0.0056	0.0010	Palm Bay, FL	-0.0003	0.0032
Akron, OH	-0.0106	-0.0060	Waterloo, IA	-0.0055	-0.0056	Miami, FL	-0.0003	0.0024
Las Vegas, NV	-0.0105	-0.0045	Lynchburg, VA	-0.0053	-0.0022	Manchester, NH	-0.0003	-0.0022
Eau Claire, WI	-0.0103	-0.0104	Omaha, NE	-0.0053	-0.0065	Eugene, OR	-0.0001	0.0021
Green Bay, WI	-0.0103	-0.0063	Tallahassee, FL	-0.0052	-0.0024	Urban Honolulu, HI	0.0001	-0.0014
Kalamazoo, MI	-0.0103	-0.0028	Durham, NC	-0.0051	-0.0037	Bridgeport, CT	0.0001	-0.0016
Wichita, KS	-0.0102	-0.0056	Cape Coral, FL	-0.0050	-0.0012	Sacramento, CA	0.0003	-0.0015
Bloomington, IL	-0.0100	-0.0094	Milwaukee, WI	-0.0050	0.0007	Trenton, NJ	0.0005	0.0009
Detroit, MI	-0.0099	-0.0011	Modesto, CA	-0.0050	-0.0047	Portland, ME	0.0007	-0.0023
Indianapolis, IN	-0.0098	-0.0038	Charlotte, NC	-0.0049	-0.0004	Billings, MT	0.0009	-0.0034
Springfield, MO	-0.0098	-0.0073	Lafayette, LA	-0.0047	-0.0088	Philadelphia, PA	0.0009	0.0014
Rockford, IL	-0.0098	-0.0025	Nashville, TN	-0.0047	-0.0028	Albany, NY	0.0010	-0.0030
Lansing, MI	-0.0097	-0.0023	Allentown, PA	-0.0046	-0.0042	Providence, RI	0.0013	-0.0006
Racine, WI	-0.0097	-0.0023	Athens, GA	-0.0046	-0.0026	State College, PA	0.0013	-0.0037
Columbia, SC	-0.0095	-0.0054	Reno, NV	-0.0044	0.0007	Bellingham, WA	0.0015	0.0009
Tuscaloosa, AL	-0.0095	-0.0124	Pensacola, FL	-0.0044	-0.0057	Charleston, SC	0.0015	-0.0030
Bowling Green, KY	-0.0092	-0.0081	Provo, UT	-0.0044	-0.0092	Burlington, VT	0.0018	0.0002
Columbus, GA	-0.0091	-0.0077	Columbus, OH	-0.0044	-0.0033	Duluth, MN	0.0018	0.0004
Greenville, SC	-0.0091	-0.0048	Albuquerque, NM	-0.0043	0.0004	Denver, CO	0.0018	-0.0009
Beaumont, TX	-0.0091	-0.0075	Colorado Springs, CO	-0.0042	-0.0042	Vallejo, CA	0.0019	0.0016
Lubbock, TX	-0.0090	-0.0126	Mankato, MN	-0.0042	-0.0036	Anchorage, AK	0.0020	0.0020
Augusta, GA	-0.0089	-0.0043	Syracuse, NY	-0.0041	-0.0035	New York, NY	0.0021	0.0003
Chattanooga, TN	-0.0087	-0.0063	Jacksonville, FL	-0.0040	-0.0015	Austin, TX	0.0025	-0.0023
Cleveland, OH	-0.0087	-0.0024	Savannah, GA	-0.0039	-0.0045	Baltimore, MD	0.0026	-0.0027
Little Rock, AR	-0.0086	-0.0052	St. Louis, MO	-0.0039	-0.0014	Virginia Beach, VA	0.0026	-0.0001
Columbia, MO	-0.0084	-0.0071	Houston, TX	-0.0038	-0.0048	Salt Lake City, UT	0.0028	-0.0009
Tulsa, OK	-0.0083	-0.0075	New Orleans, LA	-0.0037	-0.0042	Portland, OR	0.0035	0.0003
Canton, OH	-0.0083	-0.0029	Greeley, CO	-0.0036	-0.0108	Washington, DC	0.0042	-0.0021
Peoria, IL	-0.0081	-0.0069	Dallas, TX	-0.0036	-0.0025	Bend, OR	0.0042	0.0023
Cedar Rapids, IA	-0.0079	-0.0086	Buffalo, NY	-0.0035	-0.0036	Boston, MA	0.0043	-0.0019
Kansas City, MO	-0.0077	-0.0060	Deltona, FL	-0.0035	0.0002	Boulder, CO	0.0047	0.0006
Huntington, WV	-0.0077	-0.0073	Sioux Falls, SD	-0.0035	-0.0062	Charlottesville, VA	0.0047	0.0005
Lakeland, FL	-0.0073	-0.0050	Wilmington, NC	-0.0035	-0.0018	Los Angeles, CA	0.0055	0.0034
El Paso, TX	-0.0073	-0.0063	Harrisburg, PA	-0.0033	-0.0031	San Diego, CA	0.0055	-0.0012
Atlanta, GA	-0.0073	-0.0018	Roanoke, VA	-0.0033	-0.0003	Oxnard, CA	0.0055	0.0049
Bakersfield, CA	-0.0072	-0.0070	Rochester, MN	-0.0033	-0.0049	Santa Rosa, CA	0.0062	0.0033
Lexington, KY	-0.0071	-0.0052	Hartford, CT	-0.0031	-0.0024	Seattle, WA	0.0063	-0.0017
Birmingham, AL	-0.0071	-0.0063	San Antonio, TX	-0.0030	-0.0049	San Jose, CA	0.0071	-0.0031
Knoxville, TN	-0.0071	-0.0064	Springfield, MA	-0.0029	-0.0012	San Francisco, CA	0.0071	-0.0011
Huntsville, AL	-0.0070	-0.0055	North Port, FL	-0.0029	-0.0020	Midland, TX	0.0086	-0.0066